DOCUMENT RESUME

ED 304 352 SO 019 569

AUTHOR Newmann, Fred M.; And Others
TITLE Higher Order Thinking in High School Social Studies: An Analysis of Classrooms, Teachers, Students, and Leadership.
INSTITUTION National Center on Effective Secondary Schools, Madison, WI.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Jul 88
GRANT G-008690007
NOTE 222p.
PUB TYPE Reports - Research/Technical (143) -- Collected Works - General (020)

EDRS PRICE MF01/PC09 Plus Postage.
DESCRIPTORS Cognitive Development; *Cognitive Processes; Critical Thinking; High Schools; *High School Students; Research Projects; *Secondary School Curriculum; *Secondary School Teachers; *Social Studies
IDENTIFIERS *Higher Order Skills; *Thinking Skills

ABSTRACT

The papers in this report present selected findings from one phase of a five-year study on how to enhance higher order thinking skills in high school social studies classes. The study's purposes were to determine why it is difficult to emphasize these skills in social studies curriculums, what the barriers are, and how to overcome them. Five demographically diverse high schools were selected for the study. The results provided in six chapters include: (1) "The Curriculum of Thoughtful Classes" (F. Newmann); (2) "Exploring Aspects of Teachers' Thinking about Promoting Students' Thinking" (J. Onosko); (3) "Classroom Practices of High and Lower Scoring Teachers" (J. Onosko); (4) "Student Perspectives on Cognitively Challenging Curriculum" (R. Stevenson); (5) "Student Perspectives on Engaging Curriculum" (R. Stevenson); and (6) "Departmental and Principal Leadership in Promoting Higher Order Thinking" (C. McCarthy; F. Schrag). Findings indicate that classroom thoughtfulness in social studies can be assessed and can occur at high levels in conventionally organized high schools among students of all grade and achievement levels. Tables and references are included. (JHP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED304352

HIGHER ORDER THINKING IN HIGH SCHOOL SOCIAL STUDIES: AN ANALYSIS OF CLASSROOMS, TEACHERS, STUDENTS AND LEADERSHIP

Fred M. Newmann
Project Director

July - 1988

This report was prepared at the National Center on Effective Secondary Schools, School of Education, University of Wisconsin-Madison which is supported in part by a grant from the Office of Educational Research and Improvement (Grant No. G-008690007). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of this agency or the U.S. Department of Education.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

30 019 569

ACKNOWLEDGEMENTS

Papers in this volume are authored by individuals, but it should be understood that much of the planning, data collection and analysis in this study has been a collective enterprise of Cameron McCarthy, Joseph Onosko, Fred Newmann, Francis Schrag, and Robert Stevenson. In the earliest stages of planning we benefitted from the assistance of Cora Marrett, Gyu-Won Kim, Hilary McClellan and Janice Patterson, and from consultation with Catherine Cornbleth, Daniel Keating, several high school teachers in Madison, members of the Center's High School Advisory Network and National Advisory Panel. Dae-Dong Hahn provided efficient help with data processing. For secretarial services we depended on the able work of Teri Frailey, who also handled computer data entry, Sally Johnson and Diane Quayle. Larry Cuban and Oliver Moles offered insightful reactions to a draft of this report. The participants most critical to the study are the unnamed students, teachers, department chairs, and principals who were willing to share their time, their ideas, and their faith that this sort of inquiry will eventually help to enhance thoughtfulness in social studies classrooms.

TABLE OF CONTENTS

I. Introduction - Fred M. Newmann

II. The Curriculum of Thoughtful Classes - Fred M. Newmann

III. Exploring Aspects of Teachers' Thinking about Promoting Students' Thinking - Joe Onosko

IV. Classroom Practices of High and Lower Scoring Teachers - Joe Onosko

V. Student Perspectives on Cognitively Challenging Curriculum - Robert B. Stevenson

VI. Student Perspectives on Engaging Curriculum - Robert B. Stevenson

VII. Departmental and Principal Leadership in Promoting Higher Order Thinking - Cameron McCarthy and Francis Schrag

VIII. Conclusion - Fred M. Newmann

I

INTRODUCTION

Fred M. Newmann

I Outline of the Larger Study

The papers in this report present selected findings from one phase of a five-year study of how to enhance higher order thinking in high school social studies. Based on persistent claims from previous research that higher order thinking is rarely observed in high school, the larger study asks two central questions: (A) Why is it apparently so difficult to emphasize higher order thinking in the curriculum-what are the key barriers? and (B) To what extent is it possible for high school social studies departments to overcome the barriers and what is required to do so?

The full study includes syntheses of research related to the conceptualization of thinking (Newmann, in press), adolescent capacities for higher order thought (Keating, 1988), and organizational constraints on the promotion of thinking (Marrett and Kim, 1987). Empirical work will compare the degree of higher order thinking, barriers, and responses to barriers in three different sets of social studies departments: (a) those that place special emphasis on higher order thinking, but that organize instruction according to familiar patterns in the comprehensive high school; (b) those that make no special departmental-wide efforts and are conventionally organized; and (c) those that involve a departmental emphasis on higher order thinking and have also made significant changes in the organization of instruction (as suggested by the term "restructured" schools).

Selection of departments, teachers and lessons was guided by the larger project's effort to identify "exemplary" social studies departments (that is, those that make a serious departmental-wide effort to emphasize higher order thinking) and then, by contrasting these departments with others, to discover how some departments may be able to overcome barriers that others cannot. Thus, rather than concentrating primarily upon differences between individual teachers, this study explores the problem of institutionalization -what is required for departmental-wide promotion of higher order thinking?

This report focuses only on the study of select departments mentioned as type (a) above. The papers all rely upon a common conception of higher order thinking summarized in Newmann's paper, and the data for all papers was gathered according to methodology described below.

II Methodology for the Study of Select Departments

Selection of Social Studies Departments. Based on a national search, five departments from demographically diverse high schools were selected for study. Criteria emphasized a departmental-wide effort to promote thinking for all students, not only for high achievers.

Departments were selected through announcements in professional publications and special mailings to persons in the social studies community who could help to nominate schools. Several criteria were emphasized in these announcements: a departmental emphasis on higher order thinking; a required course in social studies, preferably US History, that emphasizes the study of topics in depth; the presentation of content as problematic, changing or controversial, rather than authoritative, fixed and true; students involved in problem solving and interpreting information; students reasoning about their views and receiving feedback from the teacher on the quality of their reasoning.

Sixty nominations were received, followed by correspondence, phone interviews, and a one-day site visit to nine schools. Departments in five of these schools were selected for study. To maximize candor in interviews and to protect subjects, confidentially and anonymity was promised to all individuals. Therefore, the names of schools and individuals in this report are fictitious. As illustrated in Table 1, the sample included large and medium-sized schools, schools in both single- and multi- high school districts, urban and suburban schools, and schools with a significant proportion of minority and disadvantaged students. Notably absent from this small set of schools was a small rural high school or a school with a drop out rate as large as the estimated national average of 25%.

Class Observations. The department chair at each school selected three main courses, taught by different teachers, to be observed nine times over three visits to the school in 1986-1987. The three classes were to illustrate as much higher order thinking as possible, but they were to include (a) a class with a substantial proportion of lower and middle achieving students; (b) a history course with a diverse range of students; and (c) any other class that best illustrates a concentration on higher order thinking (which usually comprised high achievers). On each visit, two additional classes, also presumed to emphasize higher order thinking and taught by different teachers, were observed once. This would result in a total of 165 lessons, but due to absences or bad weather, the final sample of lessons was 160. Within scheduling constraints, teachers were encouraged to select for our observation those lessons that placed most emphasis on higher order thinking.

Observations were recorded on seventeen 5-point scales reflecting dimensions of thoughtful discourse (e.g. Were students challenged to go beyond the information given? Were reasons carefully considered?). Descriptive notes elaborated on specific practices and responses.

Tasks assigned by the teacher that required writing were also collected and evaluated, as was written work by three students in each class.

Teacher Questionnaires and Interviews. The three teachers per school whose classes were observed nine times were designated primary teachers, and they completed three questionnaires and at least six hours of interview dealing with classroom observation and their views on a variety of issues related to the promotion of thinking. Teachers who were observed only once constituted the secondary sample, and they completed two questionnaires and one hour of interview. Total teachers = 45 (15 primary, 30 secondary).

Students. Three students from each of the 15 primary classes completed one hour interviews during each of three school visits and an total of two questionnaires. Teachers selected students representing the highest, middle, and lowest 1/3 of achievement in the class. Student data focused on engagement in schoolwork, their reactions to observed lessons and to higher order thinking tasks. Total students = 45.

The social studies department chair and principal completed questionnaires and a minimum of three one-hour interviews dealing with how the department and school organization promotes higher order thinking.

III Overview of the Papers

This study assumes that instead of trying to discover discrete teaching techniques to teach particular thinking skills to individual students, research in high school social studies should first take a broader perspective and examine how to promote more general qualities of thoughtfulness in classroom discourse. This research entails developing a method for assessing classroom thoughtfulness and also studying the views of teachers, administrators and students to learn more about barriers to its promotion and how they might be overcome.

Newmann develops the rationale for the study of classroom thoughtfulness and explains the observation scheme of 5-point scales. After presenting findings on the amount of variation within and between departments, he describes teacher and student behaviors associated with lessons that score high versus low on thoughtfulness. He concludes by examining the possible effect on classroom thoughtfulness of students' age, school achievement level and minority status.

Newmann's analysis of the lessons as a whole is complemented by Onosko's study of teachers. In what ways might teachers' beliefs about teaching influence the degree of thoughtfulness promoted in their lessons? Onosko compares the beliefs of five teachers whose lessons are consistently rated high on thoughtfulness with five teachers whose lessons receive lower ratings. He describes how the teachers differ in the instructional goals they emphasize, the complexity of their rationales, their views on the issue of breadth vs depth in covering

subject matter, their perceptions of barriers to promoting higher order thinking, and their perceptions factors that affect high and low achieving students' thinking, including the teacher's sense of influence over these factors. The second paper by Onosko describes specific ways in which higher and lower scoring teachers differ in their classroom practices and the extent to which these differences persist across all lessons.

Do students pose a major barrier to higher order thinking in school by preferring low level cognitive work? Stevenson's first paper describes the kinds of instructional activities that students find cognitively challenging. He discusses student perceptions of social studies in relation to other subjects and the kinds of tasks and mental processes that they consider most difficult, for example, their reactions to memorization, conceptual learning, making inferences and taking positions. The second paper by Stevenson describes activities that students consider engaging; that is, those that they are willing to work hard at and that they find interesting and worthwhile. Together, Stevenson's two papers offer information on the extent to which students find social studies in these select departments both engaging and mentally challenging.

While each of the departments was initially selected as exemplary for its apparent emphasis on higher order thinking, we found considerable variation among the five schools, with three high performing departments clearly separated from two lower scoring ones. How might the differences between schools explained? In studying interviews of department chairs and principals, McCarthy and Schrag examine the extent of programmatic differences, the level of resources, the nature of collegial relations, and most importantly the qualities of departmental and principal leadership that may affect the departmental level of classroom thoughtfulness. Analysis of this sort may lead ultimately to recommendations for school and district policies to minimize the barriers and maximize opportunities for higher order thinking.

This report completes some initial steps in a longer journey, but several tasks lie ahead before it will be possible to answer the main questions of the larger study. For this reason, the conclusion offers only a brief summary of the major findings of this phase in the research and an outline of analyses anticipated in the future.

References

Keating, D. P. (1988). Higher order thinking in high school: Cognitive developmental and educational constraints. Contact author, Ontario Institute for Studies in Education, 252 Bloor St. W., Toronto, Ontario, Canada, M5S 1 V6.

Marrett, C., & Kim, G. (1987). The organizational context of higher order thinking. (Final Report). Madison, WI: National Center on Effective Secondary Schools.

Newmann, F. M. (in press). Higher order thinking in the teaching of social studies: Connections between theory and practice. In D. Perkins, J. Segal, & J. Voss (Eds.), Informal reasoning and education. Hillsdale, NJ: Lawrence Erlbaum Associates.

Table 1

School Demographic Profiles

	SCHOOL*				
	A	B	C	G	S
1. 1986 Enrollment	1163	2140	2830	2022	1089
2. Ethnic Racial Composition					
i % White	65.9	75	49.9	91	32
ii % Black	2.8	9	22.2	3	67
iii % Asian	14.1	12	20.4	3	1
iv % Hispanic	12.6	4	7.3	3	0
v % Other Minority	4.6	0	.1	0	0
3. % Low Income	3	7	16	<1	42
4. Number of Teachers	72	176	145	185.6[1]	91
5. % 1986 Graduates Going To:					
i 4 Year College	40	72	60	69	37
ii Technical School	15	2	1.5	4	14
iii 2 Year Community College	30	6	30.7	5	17
iv Military	3	2	1.9	1	10
v Job + Military	12	18	3.8	14	22
6. 1986 Percent Drop-out Rate, Based on 4 Years	3.39	9.3	16.8	1	19
7. Per Pupil Expenditure 1986	3395	4100	4587	4600	5839

*Fictional names of the schools; Arnold, Bradley, Carlsberg, Grandville, Scarborough

[1]Includes specialists

II

The Curriculum of Thoughtful Classes

Fred M. Newmann

To understand the difficulties and possibilities of promoting higher order thinking, it is necessary to build both a conceptual foundation and empirical procedures that allow us to describe the extent to which higher order thinking occurs in classrooms. Relying on the conception of higher order thinking explained by Newmann (in press), this paper presents a rationale for the assessment of classroom thoughtfulness, an instrument and procedures for assessing it in classrooms, and findings on classroom practices and characteristics of schools that may account for differences between the most and least thoughtful lessons.

I Rationale for the Assessment of Classroom Thoughtfulness

In previous work that reviews literature on the nature of thinking (Newmann, in press) we have presented a conception of higher order thinking applicable to teaching a variety of subjects. Higher order thinking is defined as a challenge that requires the person to go beyond the information given; that is, to interpret, analyze, or manipulate information, because a question to be answered or a problem to be solved cannot be resolved through the routine application of previously learned knowledge. In contrast, lower order thinking demands only routine, mechanistic application of previously acquired knowledge; for example, repetitive exercises such as listing memorized information or inserting numbers into mathematical formulas. Tasks that call for non-routine mental work have been considered a distinguishing feature of higher level thinking in reviews of literature on the topic in psychology, cognitive science, education, and philosophy (Patterson and Smith, 1986; Resnick, 1987; Schrag, in press).

Simply presenting students with tasks that challenge them to go beyond the information given will not necessarily nurture success in meeting the tasks. What is needed to be successful with higher order challenges? We argue that in order to meet higher order challenges successfully, students need a combination of in-depth knowledge, skills in using knowledge, and, most importantly, dispositions or attitudes of thoughtfulness. These dimensions help to describe differences between people who succeed and fail in higher order tasks, and they suggest, therefore, three different components that ought to be taken into account when planning curriculum and instruction to promote thinking. Much of the current programmatic emphasis to promote better thinking concentrates mainly on the development of skills. A focus on skills that neglects knowledge of specific subject matter and cultivation of dispositions of thoughtfulness, however, is not likely to effect substantial, long-run improvement in students' thinking to meet higher order challenges.

Rather than selecting a specific conception of thinking such as critical thinking, informal reasoning, moral reasoning or divergent thinking, at this time it is more prudent to work toward a broad conception. Our research with social studies teachers indicates that calls for specific types of thinking are unlikely to generate widespread acceptance of any particular type. High school teachers are likely to perpetuate their previous concerns for a plurality of types of thinking, but even these will be grounded primarily in the teaching of their subjects. A broad conception that embraces diverse subjects is more likely to attract wide practitioner support.

Along with a broad conception of thinking, we recommend an approach to research that avoids fragmentation of knowledge. One might envision a long-term research effort to discover all the different ways of teaching specific content, skills and dispositions a-z to solve particular problems or challenges 1-n. Such balkanization of knowledge, however, would make it ever more difficult to synthesize findings usefully for practitioners.

Finally, a broad conception of thinking can strike at the heart of an underlying malady identified by many studies of schools. At best, much classroom activity fails to challenge students to use their minds in any valuable ways; at worst, much classroom activity is nonsensical or mindless. The more serious problem, therefore, is not the failure to teach some specific aspect of thinking, but the profound absence of thoughtfulness in classrooms (Cuban, 1984; Goodlad, 1984; Perrone, 1985; Powell et al, 1985; Stake and Easley, 1978). As suggested above, even programs to teach thinking skills can fail to promote thoughtfulness. A general conception of thinking is more likely to address the more basic issue of the general lack of intellectual challenge, because it will recognize a variety of ways in which thoughtfulness can be promoted.

One way to learn how to promote higher order thinking is to find teachers who do it rather consistently and to look for the ways in which they differ from those who demonstrate less success. If we can distinguish between more and less thoughtful classrooms, we can then inquire more systematically about the possible factors that influence the promotion of higher order thinking in classrooms. Do teachers who promote higher order thinking have different goals, conceptions of their subject, relations with colleagues and beliefs about students than teachers who teach in more routine ways? Do some departments offer specific kinds of leadership, training and support that lead to greater department-wide emphasis on higher order thinking? Does higher order thinking occur only in those classes with large proportions of high achieving students? A first step in answering these questions is to devise a measure or index of classroom thoughtfulness that distinguishes between classrooms that are more and less effective in promoting higher order thinking.

One way of documenting classroom thoughtfulness is to assess the kind of thinking demonstrated by individual students and then to

aggregate these measures for the whole class. But according to our definition of higher order thinking, individuals will differ on the kinds of problems they find challenging. For one person, trying to understand how to read and follow a bus schedule may require higher order thought, but for another, the same task will be routine. In this sense, higher order thinking is relative: to determine the extent to which an individual is facing a higher order challenge, one would presumably need to know something of the person's intellectual history. Further, to assess the extent to which the person actually participates in the analysis, interpretation and manipulation of information, one would want to "get inside" the person's head or experience his/her subjective state of thought. This, of course, is extremely difficult to do, especially with a whole class of students.

Because of these difficulties in assessing the thinking of individual students during their classes, we chose instead to examine each lesson as a whole. We may have little opportunity to assess students' individual mental states, but we can make assumptions about the prior knowledge of groups of students and, therefore, about the kinds of mental work that certain tasks are likely to stimulate. This knowledge allows us to estimate the extent to which students are confronted with non-routine challenges. We can also look for patterns in classroom discourse that reflect how thoughtfully students and the teacher cope with the challenges. That is, we can assess behavior that represents in-depth understanding of subject content, skills in manipulation of information, and dispositions of thoughtfulness.

This focus on the classroom, rather than the individual student, is not only an operational necessity; it also builds directly on a central finding from several studies of instruction in schools; namely the low level of discourse in classes, not the lack of thinking by individual students.

Having presented a conception of higher order thinking and explained the need for assessment at the classroom level, we must not neglect the more fundamental question: why are higher order thinking and classroom thoughtfulness desirable educational goals? There is much dispute on the nature of thinking and how to teach or to promote it, but the common rationales can be summarized as three main arguments. Thoughtfulness is necessary for people (a) to participate as responsible, empowered citizens in a democracy (Oliver & Shaver, 1966), (b) to contribute as productive workers in a technological society (Committee for Economic Development, 1985), and (c) to have rewarding personal lives which includes managing one's private affairs, continuing to learn and benefitting from culture (Schrag, in press).

We assume that thoughtful lessons, by presenting students with higher order challenges and nurturing the knowledge, skills and dispositions required to meet the challenges successfully, will enhance these benefits for individual students. This stage of the empirical research does not assess the effects of classroom thoughtfulness on individual students, but we would predict that if individual

assessments were conducted, students who participate in more thoughtful classrooms will have greater success on oral and written assignments that pose higher order challenges in subjects closely related to the class. Such students are also more likely to succeed on tests of simple information retrieval in the subjects studied. A future phase of the research will examine the relationship between classroom thoughtfulness and student achievement.

Apart from its effects on individual achievement, however, the promotion of classroom thoughtfulness can be defended as worthwhile on the ontological claim that using the mind to confront new challenges is critical to human nature (Rawls, 1971). To the extent that we are deprived of the opportunity to work on these challenges, we are unable to express a central part of our being.

Having explained the conceptual basis of the study, the ensuing report of empirical work addresses two main questions: (A) What dimensions of teacher and student discourse can be used to describe the degree of higher order thinking promoted in social studies classes? (B) What teaching practices, characteristics of schools, and characteristics of students tend to distinguish between the most and least thoughtful lessons?

II An Instrument to Assess Classroom Thoughtfulness

A. Guidelines

Several considerations guided the development of our instrument. The principal concern was to develop an observation scheme consistent with the general conception of higher order thinking outlined above. More specific guidelines such the following were also used:

> The categories or variables should be able to be observed in the teaching of a variety of subject matter and skills within social studies.

> The categories should refer to teacher behavior, to student behavior, and to activities involving both teacher and student.

> As an exploratory instrument, the categories should include many dimensions which, on the basis of further theoretical and empirical analysis, might later be reduced to a smaller number of essential criteria.

> The categories should be conceptualized in ways that might later be used to help teachers reflect on their practice.

Using these guidelines, the research team considered the theoretical and practical advantages of a large number of criteria. Based on observations of several live and videotaped classes, discussions with experienced teachers, and exercises to achieve inter-rater reliability, seventeen observational scales were selected.

B. The Scales

The scales are described in detail in the Appendix to this paper. General characteristics of the lesson which are normally the responsibility of the teacher include 1. sustained examination of a few topics rather than superficial coverage of many, 2. substantive coherence and continuity, and 3. sufficient time to think. The absence of any of these will, we presume, undermine the disposition to think things through systematically and deliberately.

Several scales focus more specifically on teacher behavior. The thoughtful teacher must 4. ask challenging questions or pose challenging tasks. In guiding students' work with the tasks, the teacher should 5. carefully consider explanations and reasons for conclusions, 6. press individual students to justify or clarify their assertions, 7. encourage students to generate original and unconventional ideas, 8. show an awareness that not all assertions emanating from authoritative sources are absolute and certain, 9. integrate students' personal experience (where relevant) into the lesson, and 10. model other characteristics of a thoughtful person.

These teacher characteristics will help to reinforce in-depth understanding and the qualities of reflection characteristic of effective thinkers. Conversely their absence undermines student success with higher order challenges. For example, a teacher who does not press students to clarify or justify their answers reinforces their tendency to say whatever comes into their heads without considering its warrant or implications. Or, a teacher who doesn't give students encouragement for legitimate but unconventional responses supports the stereotyped responses which characterize weak thinkers.

Several scales focus on student behavior. In thoughtful classrooms, students 11. offer explanations and reasons for their conclusions, 12. generate original and unconventional ideas, 13. assume the role of questioner and critic, and 14. make contributions which are articulate and germane to the topic. To behave in these ways, students must participate verbally (15, 16). If these behaviors are not evident, other cues may indicate engagement in mental work (17).

Based on criteria which will be explained later, six of the scales were eventually selected as minimal indicators of classroom thoughtfulness.[1]

---

[1] Our basis for determining thoughtfulness was verbal discourse dealing with the content of lessons, but indicators of thoughtfulness in other areas might also have been examined; for example, rules for student behavior and handling of violations; grading practices; assigning student responsibility for non-instructional tasks such as classroom clean-up and distributing materials. Thoughtful discourse in the formal part of lessons may be related to these aspects of classroom management. These areas

C. The Absence of Developmental and Hierarchical Cognitive Schema

Readers may wonder why the instrument fails to incorporate criteria that represent a progressive hierarchy from lower to higher level thought, such as a Piagetian scheme from concrete to formal operations or a Bloomian taxonomy of cognitive objectives from knowledge to evaluation. In addressing this issue Schrag (1987) has shown several reasons why such schemes are unlikely to be helpful. In spite of extensive research, writing and attempts to apply these schemes to practice, at least two serious problems remain.

First, the wording of a task alone is insufficient for determining the extent to which the student has to interpret, manipulate or evaluate information in new or non-algorithmic ways. Consider the test question, "Which of the following would be the likely result of maintaining a high protective tariff? (a) higher prices for domestic goods; (b) lower prices for foreign goods; (c) increased foreign trade; (d) higher prices for farm products sold in foreign markets." If the results of protective tariffs were described clearly in the text or class, this task would involve only recall, but if only the definition of a high protective tariff were given, with no explanation of the historical and economic reasoning behind it, it would require students to make original inferences. That is, the cognitive demands of a task cannot be estimated without knowledge of the tools and resources, or prior experience, the student has available to meet it.

Second, in spite of advances in cognitive science and the study of human information processing, there is no consensus on the ways of categorizing cognitive operations and the extent to which some operations are inherently more complex and difficult than others. It has been shown that the apparently simple act of decoding the meaning of a word requires many cognitive operations. Sternberg's (1983) analysis of the solving of multiple-choice analogy problems revealed seven main components: encoding, inference, mapping, application, comparison, justification, and response. With this level of complexity in trying to describe how the mind works, it becomes ever more difficult to compare the levels of thinking required by different problems or subject domains. Would a sophisticated analysis of the political effects of increasing tariffs necessarily involve more or less complex thinking than analysis of economic consequences?

In short, the specific kinds of cognitive work generated by any given task or question depends largely upon the often unknown resources and knowledge of the learner, and even if the nature of this work could be described with greater precision, it is particularly difficult to

---

were not systematically observed, but no striking differences on such matters were reported by researchers.

justify the labeling of some cognitive operations as more complex or at a higher level than others.[2]

Although fine discriminations in levels of cognitive operations are difficult to make, we can still determine in a more general way whether students are being challenged to use knowledge to go beyond the information given to solve non-routine problems. That is, when we know what knowledge students are generally familiar with, what materials they have at their disposal, what kinds of responses are expected from teachers, and how students behave in response to classroom tasks and questions, our scales enable us to estimate the extent to which particular problems are intellectually challenging and the extent to which students and teachers approach the work in a thoughtful manner.

D. Rating Scales vs Categorical Frequencies

Schrag (1987) reviewed the advantages and disadvantages of using rating scales and category schemes to record classroom discourse. We chose rating scales for three main reasons.

First, and most important, categorizing and counting highly specific behaviors can yield misleading indicators of the quality of discourse; rating scales allow more sensitivity to the context in which behavior occurs. For example, one teacher might ask 20 questions that would be considered "challenging" in the sense of requiring students to go beyond the information given. But this teacher might give students most of the answers after only superficial discussion. In contrast, another teacher might pose only one challenging question at the beginning of class, but spend the entire class period prodding students to develop increasingly sophisticated responses. The second classroom, which had a much lower frequency of challenging questions would be more substantively challenging than the first.

To further illustrate the problem, consider scale #5, "The teacher carefully considered reasons and conclusions." In making estimates about the nature and extent of this behavior, and other forms of teacher-student interaction, it is often necessary to take into account three considerations simultaneously, each of which could technically be counted: the amount of time spent on the behavior; the number of discrete occurrences of the behavior; and the number of students involved. Such a recording scheme would be operationally cumbersome (if not impossible), and as indicated above the quantitative sum of these three dimensions could still give a misleading indication of quality.

Second, some of the dimensions of interest are conceptually grounded in qualitative criteria that are likely to be violated by

---

[2] Further support for our concerns with the limitations of cognitive developmental schema is provided by Keating (1988, in press) and Keating and McLean (1987).

translation into a frequency count; for example, "The lesson displayed substantive coherence and continuity," "The teacher was a model of thoughtfulness." For items such as these, it is difficult to imagine what might be counted that might still capture the intent of the criteria.

Finally, as indicated under guidelines, we wanted to generate criteria that might eventually be used by teachers to reflect on their own teaching. Categorical frequency systems are likely to become too complex and elaborate for teachers to use, but a set of scales can be adapted to a more practical observation scheme.

Using the scales to guide assessment of thoughtfulness in classrooms, we studied the teaching of social studies in five high schools.

III Methodology

The selection of departments and classes to be observed was described in the introduction. In each of the schools, three teachers were observed over nine classes, and six teachers once. The classes were selected to include students of diverse levels of school achievement.

A. Reliability

Drawing from a team of four researchers, two members gathered data at each school, and the two-person teams varied across schools. During each visit they jointly observed one class from each of the three main teachers. Ratings were made independently on each of the scales. Following the lesson, discrepancies between ratings were discussed. To estimate inter-rater reliability, Pearsonian correlations were computed between the original ratings of 17 variables by each researcher on each lesson. For each visit the correlation between 51 ratings (3 lessons x 17 scales) was also computed. As shown in Table 1, inter-rater agreement was well above .80 for most comparisons, and improvements occurred especially between visits 1 and 2. The degree of rater agreement for each scale across all visits is indicated in Table 2 (see Table 3 for scale names and content). For fifteen of the seventeen scales raters agreed precisely more than 50% of the time. If the criterion is expanded to include a difference of no more than one point, we find that all of the seventeen scales attracted rater agreement at least 90% of the time.

B. Selection of Minimal Criteria for Classroom Thoughtfulness

The exploratory nature of the study led to creation of a large number of indices of thoughtfulness, but it would be useful to reduce these to a smaller number of essential variables. First, if the long list can be summarized into a more parsimonious set, eventually it will be more feasible for teachers to use in reflecting upon and guiding their teaching. Second, a smaller set can be helpful in examining the

validity issue. That is, if the scores of lessons on the minimal criteria also correlate with other criteria of thoughtfulness, we can have more confidence that we are assessing an important underlying feature of classroom discourse.

Presumably, each of the scales represents a desirable characteristic that would contribute to thoughtful discourse. Nevertheless, it could be useful to make a distinction between a criterion that indicates or helps to promote higher order thinking versus one that, in addition, seems so essential that one could not imagine judging a lesson "thoughtful" unless the criterion were met. Since we were not able to find analytic or empirical literature that conclusively determined a few key criteria, it was necessary to rely largely on deliberations within the research team.

In deliberating on how to select the most essential criteria, we put each scale to the following test: Based on the conception of higher order thinking outlined earlier, could a lesson conceivably score low on this scale, yet still be considered a highly thoughtful lesson? If the answer was "yes," then the scale was not considered critical as a minimal criterion. If the answer was "no," the scale was judged as being minimally necessary, though perhaps not a sufficient, criterion for thoughtfulness.

Without explaining our decision for each scale, a few examples will illustrate the reasoning used to disqualify four of the scales as minimal criteria. Teachers may conduct challenging large-group discussions in which they ask provocative questions, orchestrate student responses to each other, and offer important information, but they may rarely interrogate individual students through Socratic questioning that involves several teacher-student exchanges (#7). A lesson might carefully consider the reliability of primary sources in describing events that triggered the American Revolution but not necessarily result in students' offering original and unconventional solutions to the problem (#13). A thoughtful discussion of the conflict between freedom of speech and national security might focus on the quality of reasoning found in two editorials without the teacher's integrating students' personal experience into the lesson (#4) and with relatively few students participating (#15).

On the other hand, it would be impossible to give recognition for thoughtfulness if the lesson failed to illustrate our defining feature of higher order thinking, that is, tasks or questions that pose cognitive challenge (#6) to go beyond the information given. Similarly, it would be inappropriate to give a high rating to a lesson that lacked substantive coherence about the subject of study (#2), or one in which students gave no reasons or explanations for their views (#12). As a result of these kinds of considerations, we selected the following six minimal criteria:

Sustained examination of a few topics rather than superficial coverage of many. (1)

Substantive coherence and continuity. (2)

Students were given an appropriate amount of time to think and to respond. (3)

The teacher asked challenging questions or structured challenging tasks. (6)

The teacher was a model of thoughtfulness. (10)

Students offered explanations and reasons for their conclusions. (12)

These criteria eventually formed the basis for assessing the degree of thoughtfulness demonstrated in each of the 160 lessons. A case might be made for including other criteria; for example, teacher careful consideration of reasons (#5), or students as questioners and critics (#11). Our intent here is not to settle the question of essential criteria for classroom thoughtfulness, but only to develop one reasonable, manageable set that allows us to explore the degree of variation among classes and schools, along with possible explanations for that variation.[3]

Ratings on the minimal criteria permitted us to identify the most and least thoughtful lessons as well as the teachers and the departments whose lessons were rated consistently high or low. Having identified the essential criteria, the strategy was to search for other characteristics of lessons, teachers and schools that might be associated with these indicators and that might, therefore, suggest "causes" or interventions that would help teachers and departments move from lower to higher levels.

IV Findings

A. Frequencies and Relationships Among Dimensions of Thoughtfulness

The schools were selected for their emphasis on higher order thinking, but the selection process did not allow for data collection to determine levels of thoughtfulness according the scales. We were first interested, therefore, in the frequencies of ratings on all the scales across all lessons. This information is given in Table 3. Some dimensions occur very rarely, especially Socratic questioning (#7), integrating student experiences into the lesson (#4), teachers trying to get students to generate original ideas (#8), teachers questioning authoritative sources (#9), and students engaging in thoughtful

---

[3] Later it will be shown that many of the scales correlate highly with one another. It is likely, therefore, that certain modifications in the set of minimal criteria would not significantly alter findings based the original set.

discourse with other students (#16). Evidently even these select social studies departments neglect these aspects of thoughtfulness.

On the other hand, we found a relatively high occurrence of each of the six minimal criteria and other dimensions as well; for example students assuming the role of questioner and critic, student contributions being articulate and germane, student verbal participation and involvement. If these departments are truly exemplary, higher frequencies on these characteristics would presumably distinguish them from a more representative sample of schools, and this will be tested in future research.

To what extent are the aspects of thoughtfulness related? If a lesson shows a high rating on cognitive challenge, is it likely also to show a high rating on other dimensions? Table 4 presents correlations among the 17 scales. Generally, we find a high degree of interrelationship among those scales that show variability across the five (or four) scale points. In contrast, the highly skewed dimensions of enough time to respond (#3), integration of student experiences (#4), teacher awareness of controversy in authoritative sources (#9), and students engaged in discourse with one another (#16) show little relationship to the other scales. If these low variance items were removed from the correlation matrix, 62 of the 78 remaining correlations (80%) would be significant (p .05). This supports the possibility that the 13 remaining scales may be measuring a single underlying variable. Later, in presenting findings on the differences between lessons that score high versus low on the minimal criteria, it will be shown that these two groups of lessons also vary in the expected direction on several of the other criteria as well, which lends further support to the existence of an underlying quality.

If these variables do represent an underlying construct of thoughtfulness, it is also possible that thoughtfulness itself may be composed of slightly distinct dimensions. We explored this briefly through factor analysis of the 17 scales. This suggested two main components, one grounded primarily in teacher behavior and the other in student behavior. Table 5 gives the results of a 3-factor model in which 8 of the variables loaded highly on a student participation factor (#1 which explained 24% of the variance), and 6 of the variables loaded highly on a teacher thoughtfulness factor (#2 which explained 23% of the variance). The third factor was not interpretable. The main factors might represent two different teaching styles: one in which the teacher takes a highly directive role in guiding rigorous consideration of challenging topics, and another characterized largely by idea-swapping or sharing among students with the teacher orchestrating, but giving less substantive direction to the discussion. Further analysis of relationships among the scales will be addressed in the future.

B. Accounting for Differences Between More and Less Thoughtful Lessons

The ultimate purpose of constructing quantitative indicators of classroom thoughtfulness is to learn more about how to promote higher order thinking in the classroom. To pursue this we examined the potential influence of other variables that may affect the degree of thoughtfulness as measured by the six minimal criteria. We estimated the relationship between levels of minimal thoughtfulness and two general sets of variables: classroom behavior and practices that teachers can alter; and institutional and background variables such as the influence of schools vs individual teachers, and students' age, school achievement level, and race-ethnicity.

The first step in this process was to create two groups of lessons that represent, respectively, high versus low levels of higher order thinking. To do this, we averaged each lesson's ratings on the six minimal criteria (HOTAV) and then selected for comparison those in the top and bottom quintiles.[4] Based on this procedure, the high group included 36 lessons with a mean HOTAV of 4.60 (sd .21) and a low group of 42 lessons with a mean HOTAV of 2.78 (sd .23). The difference between the two groups is substantial, representing 2.68 standard deviations of the full group of 160 lessons. Although the lessons were drawn from social studies departments that all emphasized higher order thinking, we see here considerable variation in the degree of classroom thoughtfulness among lessons.

1. Two lessons.

To give a sense of the variation in actual classes, we describe below a lesson from each group.

---

[4] We anticipated the possibility that average scores on the minimal criteria might lead to inappropriate selections (e.g. a lesson with a 1 on teacher challenge, and a 2 coherence, but with 5's on the four other variables shouldn't be considered high). Initially, therefore, we required specific levels for each criterion. "High" lessons were those that scored 5 on time and greater than or equal to 4 on few topics, substantive coherence, teacher challenge, teacher model, student explanations. "Low" lessons were initially defined as those that scored less than or equal to 4 on time, and less than or equal to 3 on the five other criteria. This procedure resulted in selection of 18 (11% of the total) high lessons and 12 (8% of the total) low ones, but several lessons in each group were taught by just a few teachers. Since we wanted to analyze a larger proportion of the lessons and include a greater variety of teachers, we changed the selection criteria to the HOTAV method. In comparing these two selection procedures, we found that the averaging method did not result in inappropriate selections originally anticipated, and the pattern of differences between the high and low groups other variables remained the same.

Lesson 1

In this US History class, the main topic was the nullification controversy between South Carolina and President Jackson when the state declared the tariffs of 1828 and 1832 to be null and void. Students had read textbook material on the background, but also excerpts from three source documents: South Carolina's "Address to the People of the United States," November, 1832 (anonymously written by Vice President John C. Calhoun); Andrew Jackson's "Nullification Proclamation," December, 1932; and Calhoun's speech to the Senate, February 15-16, 1833, in defense of nullification and in opposition to the Force Bill then being considered in Congress.

The teacher began by asking for students' reactions to the source readings. "Why are they more difficult?" she asked. Students had trouble articulating reasons, and she explained that sources may seem more difficult than textbook reading, because, unlike textbooks, the political speeches were intended not to create an understandable story for 20th century students, but to speak directly to the people and issues of former times.

The teacher then asked for definitions of key terms. "Nullification," and "tariff" were quickly defined, but it took a long series of patient questions to develop an understanding of why the tariff would benefit northern manufacturing states to the disadvantage of southern agricultural ones. During the discussion, students frequently spoke up when they failed to understand, and the teacher tried to clarify the economic relationships by using the contemporary analogy of the effects on various interests of an import tax on Japanese cars.

Sensing that the impact of the tariff was reasonably well understood, the teacher then asked students to examine the "Address to People of the United States" and to pick out the main arguments for nullification. One student summarized the main point that since the states had the power to form the Constitution, they also enjoyed the authority to nullify any actions of the Federal government not specifically delegated to it. Another student noted the point that the states' ratification of the Constitution included a special obligation to protect it from usurpation of power by the Federal government. The teacher pressed students to find further arguments, and they did: the claim that the Federal government may tax only to raise revenue, but not to protect some internal interests to the detriment of others; and the argument that Jackson's threat to use force on this issue would supercede the law. In each instance, the teacher focused on wording in the text and asked students to give their own understanding of the argument.

In a similar way students searched for Jackson's arguments: nullification itself would be unconstitutional; that the states, by joining the Union, had given up the right to secede; that his election

by all the people, gave the Federal government powers beyond the original power of states to enforce Federal actions.

When the alternative arguments seemed clearly understood, the teacher asked students whether they could think of any conceivable situations in which they felt a state might be justified in nullifying a Federal action. They had some difficulty in proposing specific, credible policies, but they did suggest such situations as the Federal government wanting to build a nuclear base in a state; or Congress making smoking illegal (which would be opposed by tobacco producing state). Discussion of hypothetical situations failed to produce a conclusion on when nullification might be justified, but the teacher alerted students that the issue will come up again.

This lesson received high ratings for focusing on a few topics, substantive coherence, enough time to respond, careful consideration of reasons and explanations, Socratic questioning, teacher modeling thoughtfulness, students' giving explanations and reasons, students generating original ideas, student comments being articulate and germane. More than 75% of the students participated, and between 50% and 75% showed genuine involvement. The HOTAV score for the lesson was 4.5.

Lesson 2

An Advanced Placement US History class considered the South's resistance to the abolition of slavery. The teacher began by outlining why three approaches to the abolition of slavery each seemed unfeasible prior to the civil war: sending slaves back to Africa; gradual emancipation with financial compensation to slave owners; and immediate legal abolition. None of these were discussed in much detail, but after the teacher indicated that in the long-run, slavery was not economically feasible for the South, students began asking a number of questions: How much did slaves cost? Why didn't the South rely on cheaper immigrant labor which would also give them more representatives in Congress? What was happening in the cotton business? Didn't the South see they were being left behind (in economic growth) by the North?

The teacher responded to the questions with a set of short lectures on the Southern cultural attachment to the institution of slavery, the relationship of British trade to the economy, and the resources (human, capital, technical) that would be needed for the coming industrial revolution.

After recognizing these as slight digressions, the teacher returned to describe the abolitionists and their movement. She described their writing, and emphasized how they used the media to influence public opinion. When she pointed out that even the abolitionists did not support full equality for blacks, a student asked when fuller equality was achieved. In another short lecture, the

teacher reviewed the long history of gaining legal rights for blacks through the 1960's.

A student remarked that it took the US at least 100 years to develop equal rights in the law, and that it seems that some people expect South Africa to do this overnight. This led to teacher and student comments on parallels and differences between the US and South Africa. Students also referred to British resistance to colonial independence (especially India). A wide-ranging, unfocused discussion ensued on imperialism and foreign policy issues related to Ireland, Israel, Iran, and Nicaragua. Although the common theme seemed to be importance of economic self interest versus some higher morality, the discussion was not directed to examine this systematically.

In this lesson, students consistently asked provocative questions, they generated original ideas, and they were articulate. More than 75% participated verbally, but less than half showed genuine involvement, and usually students failed to explain or give reasons for their ideas. The teacher, while responding in an academically respectable way to student inquiries, did not pursue any topic in depth, nor did she pose challenging tasks for the student. She did not press for careful consideration of her own or the students' reasons, and she indicated more of a sense that she had the answers than a spirit of reflectiveness in working through a problem. The HOTAV score for this lesson was 2.5.

2. General Differences between High and Low Lessons.

Moving from these examples, to the mean scores for the high and low lessons, we consider now how the high and low lessons differ on variables other than the minimal criteria for thoughtfulness. Table 6 shows comparisons between the high and low groups on each of the scales.[5] To gain a sense of the relative size of differences between

---

[5] The entire pool of lessons included nine lesson observations from 15 teachers and one lesson observation from 30 teachers. To maximize the number of cases, we used the entire pool of 160 lessons for quantitative analyses. To determine whether results would be biased by overrepresentation of lessons from the 15 primary teachers, we conducted parallel analyses using only one set of lesson ratings from each of the 45 teachers. For this purpose, primary teachers' ratings across their nine lessons were averaged to create a single lesson profile. These "unweighted" analyses yielded findings consistent with those in Tables 6, 9, 12, and 13, and we concluded that use of the entire pool of lessons produced acceptable estimates.

groups, the table expresses the differences between means in terms of the standard deviation of all lessons.[6]

Of the 11 non-minimal criteria, the high and low lessons seem to differ most on five dimensions: teacher careful consideration of student reasons (#5), teacher Socratic questioning (#7), students contributing original ideas (#13), students being articulate and germane (#14), and students being involved in the lesson (#17). For each of these dimensions the mean difference between the groups exceeded 1 point on a five-point scale, and the effect size for each also exceeded one standard deviation.

While some differences appear on the other six scales, it would be difficult to make a case that any of them are substantial, for no others reach a one standard deviation degree of difference.[7] A number of interpretations might be made of these results, but we wish to highlight three. First, only two criteria with clear implications for teacher behavior stand out: teacher careful consideration of the quality of student reasons, and Socratic questioning. We included in the scheme three other teacher behaviors that might theoretically might appear to promote thoughtfulness - integrating student experiences into the lesson, encouraging students to contribute original ideas, and showing awareness that not all authoritative sources provide adequate answers. None of these, however, distinguished between the high and low lesson groups.

---

[6] Results for analysis of variance are given in Table 6 for the reader's information, but not discussed, because it is problematic as to whether the set of lessons can be considered a random sample of a larger population. In a technical sense, lessons were not randomly sampled, either from the full universe of high school social studies lessons, or from all the lessons in the select departments. On the other hand, one might assume that the select departments do represent a larger universe of departments emphasizing higher order thinking. Furthermore, since the observed lessons involved a diverse range of students and courses, occurred at several points in time, and were selected primarily according to scheduling convenience, rather than teaching technique or topic, they might be assumed to represent the larger universe of lessons in the departments. In any case, with this small sample we believe the size of the differences is a more meaningful indicator than the level of statistical significance.

[7] The comparisons between the high and low groups are corroborated by the Pearsonian correlations of the other scales with HOTAV as shown in Table 4. Proportion of students who talk (#15, r=.34) appears a little stronger here, but this may be largely a result of its own correlation with articulateness and germaneness which is .52.

Second, neither the amount of student questioning and criticism (#11), nor the proportion of students who participate (#15), differed substantially between the high and low groups. This supports the point that the quality of student participation is probably a better indication of classroom thoughtfulness than the quantity. This is further reinforced by the fact that indicators more likely to tap quality, namely, student originality and student articulateness and germaneness, did differentiate in the expected direction.[8]

Third, given a common observation that many students may resist higher order thinking because they are presumed to prefer more structured routines of school that may seem easier to master, it is encouraging to find here that in the more thoughtful lessons a higher proportion of students consistently showed signs of engagement (#17) with the material.

Did the high lessons differ systematically from the low in certain generic teaching practices such as the use of lecture, the type of reading assigned, or the kinds of writing done? Each lesson was categorized according to the presence or absence of certain practices. We asked which of the following was the dominant practice during the lesson: lecture, film, recitation, teacher-centered discussion, student centered discussion, student groupwork, individual student seatwork, oral reports, and other. Types of reading were classified as textbooks, articles, primary sources, literature, other and none. Writing done in class was categorized according to outlining, notes, worksheets (complete sentences required), complete sentence answers, paragraph or essay, research paper, other. Each of the reading and writing items received a score of 1 if it occurred during the lesson, 0 if it did not.

Table 7 shows the results of comparison between the high and low groups. Differences appeared in only a few areas. The dominant practice in the more thoughtful lessons was teacher-centered discussion, but the less thoughtful lessons were dominated by lecture, recitation, and they also included teacher-centered discussion. More thoughtful lessons relied less on textbooks and more on primary sources and other types of reading. The less thoughtful lessons relied largely on a textbook.

The findings confirm what might be expected: in contrast to lecture and recitation, teacher-centered discussions are more likely to challenge students to go beyond the information given, as is the reading of primary sources, in contrast to textbook reading. One might expect the more thoughtful classes to involve more extensive prose

---

[8] The proportion of time students were engaged in thoughtful discussion with one another (#16) might be considered a combination of quantity and quality. Its failure differentiate between the groups can probably be attributed to its extremely rare occurrence (it had the lowest mean of all student behavior items).

writing, but in neither group of classes did we observe much writing other than note-taking and completion of short answer exercises.

There were no substantial differences in the forms of writing done between the more and less thoughtful lessons, but we found some differences in the quality of written tasks. Written assignments distributed during our observations were collected, and the tasks were coded according to the extent to which they required students to (a) draw inferences, (b) give reasons, (c) integrate information from a number of sources, (d) develop an idea or theme; also, whether they allowed students to generate original responses, and whether students had an appropriate amount of time to complete the task. Ratings on each task were added to produce a higher order challenge score. Although the scores of all tasks were rather low (less than 2 on a 3-point scale), tasks assigned in the top quintile of lessons did exceed those in the bottom quintile by about .75 of bottom group's standard deviation.

3. The influence of institutional and background factors on classroom thoughtfulness.

Having found some aspects of teacher and student behavior associated with differences between the high and low classes, it is important to ask about the influence of institutional and background factors. Is it possible, for example, that even with considerable variation among teachers, some social studies departments have consistently more thoughtful lessons than others? Would teachers assume that older students are more capable of higher order thinking, and therefore, emphasize higher order thinking more frequently in 12th grade than 9th grade classes? Would classes with a preponderance of high achieving students reflect consistently more thoughtfulness than classes with large proportions of low achievers? Would the proportion of minority students in a class tend to decrease the level of thoughtfulness? Are elective courses more likely to promote thoughtfulness than required courses? We examined these issues by regressing HOTAV on the relevant variables.

a. The Impact of School.

The school means and standard deviations on HOTAV given in Table 8 show that the five schools do differ. Indeed, they can reasonably be seen as two groups, with Carlsberg, Grandville, and Arnold distinguishing themselves from Bradley and Scarborough. Regression analysis (Table 9) indicated that differences between schools accounted for about 26% (adjusted r square) of the total variance among all lessons. While most of the variance in individual lessons is due to other factors, even in this group of presumably exemplary departments, a significant portion of the variance can be attributed to departmental (or school) differences. The paper by McCarthy and Schrag explores how these differences might be explained by different patterns of institutional support.

b. Impact of the class grade level, students' average school achievement level, percent minority, and required or elective status.

For each lesson, teachers indicated the percentage of students in enrolled in the 9th, 10th, 11th, and 12th grades; the percentage of students whose grade point average placed them in the lowest, middle and highest one-third of the school achievement distribution; the percent of minority students (black, hispanic, asian, other); and whether the class was required or elective. The Pearsonian correlations between HOTAV and each of these variables is given in Table 10 .

The proportions of students in the different grade levels are inconsistently associated with classroom thoughtfulness. Although there may appear to be an advantage in having lots of 10th graders and few 11th graders, these relationships are not robust. The proportion of students in the lowest, middle, and highest thirds of school achievement has no relationship to higher order thinking, which helps to establish the prospect for promoting thoughtfulness equitably for students of all achievement levels. The results on minority composition of the class indicates less higher order thinking with larger proportions of blacks, and more with larger proportions of hispanics, but these results should be interpreted in the light of the very small proportions of all minorities other than blacks and their overall distributions in the schools as shown in Table 1 (Introduction) and in the observed lessons as shown in Table 11. In the observed lessons, the maximum percent for hispanics, for example was 15% or 4 students in a class of 25.

To estimate the impact of these background factors independent of the variance they share with one another, they were regressed together. To simplify the analysis, composites were created for grade level and achievement level, and the minority concentration variable was restricted to blacks. Table 12 indicates that together, these variables account for only a small amount of the variance in classroom thoughtfulness (adjusted r square = .11), and that this is due largely to the slight negative relationship associated with the proportion of black students in a class. Because of the small magnitude of this finding, we should not make too much of it, but it does suggest that independently of students' achievement level, teachers expose classes with large proportions of blacks to less higher order thinking.[9]

Except for this finding, the results are most encouraging, because they illustrate that three other commonly perceived obstacles to promoting higher order thinking (student youth, low school achievement, required courses) need not stand in the way. That is, in these

---

[9] On average, an increase of one standard deviation, or 24%, in the percentage of blacks in a class could be expected to decrease the lesson's HOTAV by .38 standard deviations, or by .26 points on the five point scale.

schools, chosen for their emphasis on higher order thinking with a diverse population of students, the level of classroom thoughtfulness is due to other factors, primarily, we assume, to the commitment and craft of individual teachers.

Although the background variables as a group have no impact on lesson thoughtfulness, it is possible that their influence and that of the schools might change when the institutional and background variables are considered in one regression model. For example, if some background variables such as proportion of minorities in classes or proportion of ninth graders in classes are much higher in some schools than others, one would want to examine their effect on thoughtfulness independent of the impact of the school culture. Table 13 presents this information.

After taking the school into account and comparing this information with Table 12, grade level and required course now show a possible (though minuscule) influence on thoughtfulness, favoring younger students and elective courses. In contrast, the influence of students' achievement level disappears even further, and the effect of proportion of blacks is also reduced to non-significance. By considering the impact of school and background variables together, we see also that the differences between schools have been slightly reduced (compare with Table 9). We can also probably assume that the previously noted effect of proportion of blacks was probably due to that fact that the school scoring lowest on thoughtfulness also taught classes with the highest proportion of blacks. By adding background characteristics to the model, we have explained only an additional 1.5% of the variance in lessons' higher order thinking (subtracting the adjusted r square of Table 9 from that of Table 13). This further underscores the impact of institutional (possible departmental) culture.

V Conclusion

We began by reviewing a conception of higher order thinking that emphasizes teaching the knowledge, skills and dispositions necessary to meet non-routine intellectual challenges. Next we proposed that in order to improve practice in this direction, it would be useful to assess levels of classroom thoughtfulness in ways that would distinguish between lessons that are more and less successful in the promotion of thinking.

The empirical study focused initially, therefore, on the methodological question, "What dimensions of teacher and student discourse can be used to describe the degree of higher order thinking promoted in social studies classes?" We identified seventeen possible dimensions and found that they could be reliably coded. The overall ratings in 160 lessons showed that several of these dimensions rarely occurred, but most of them manifested reasonable variance and correlated with one another in ways that would be expected from previous research and experience. Exploratory factor analysis

suggested the existence of two main teaching styles: one in which the teacher rigorously presses students to explain their ideas and one in which the teacher guides the discussions less directly by encouraging more free-wheeling student interaction.

The second main question was, "What teaching practices, characteristics of schools and characteristics of students tend to distinguish between the most and least thoughtful lessons." Six dimensions were selected as minimal criteria for classroom thoughtfulness. Comparisons of the top and bottom 20 percent of the lessons on these variables indicated that two main teacher behaviors (careful consideration of reasons and explanations; and Socratic questioning), and two main student behaviors (generation of original ideas, and articulateness and germaneness) are most associated with scores on the six essential criteria. It was also encouraging to find that students were more engaged in the more thoughtful classes. The most thoughtful lessons were dominated by teacher-centered discussion, and the least thoughtful more by lecture and recitation. The most thoughtful lessons also involved less reliance upon textbooks and more upon primary sources and other forms of reading.

The level of classroom discourse in lessons varied as a function of the school (department) in which the lesson occurred, but did not vary according to the grade level of the lesson, proportion of minority students, the general achievement level of students in the class, or whether the course was required or elective. Although, differences between schools had clear impact, most of the variance in classroom thoughtfulness was due to unmeasured factors, and the most powerful of these probably relate to teachers' individual commitments, orientations and skills. The paper by Onosko compares teachers whose lessons show most consistent emphasis on thinking with those who have less consistent success.

Since the study was confined to only five schools, chosen for their presumably exemplary emphasis on higher order thinking on a department-wide basis, these findings may have only limited generalizability. Future research will examine the applicability of the findings in two additional sets of schools: a representative group in which no school claims a departmental-wide emphasis, and a set of schools that claim to emphasize higher order thinking, but that have also made significant changes in the organization of instruction in order to fulfill this purpose. Examining levels of thoughtfulness, the influence of institutional and background factors, and differences between teachers in the three sets of schools should give us further clues as to how to promote higher order thinking in social studies.

References

Committee for Economic Development (1985). Investing in our children: Business and the public schools. New York: Committee for Economic Development.

Cuban, L. (1984). Policy and research dilemmas in the teaching of reasoning: Unplanned designs. Review of Educational Research, 54(4), 655-681.

Goodlad, J.I. (1984). A place called school: prospects for the future. New York: McGraw Hill.

Keating, D. P. (1988). Higher order thinking in high school: Cognitive developmental and educational constraints. Contact author, Ontario Institute for Studies in Education, 252 Bloor St. W., Toronto, Ontario, Canada, M5S 1 V6.

Keating, D. P. (in press). Structuralism, deconstruction, reconstruction: The limits of reasoning. In W. F. Overton (Ed.), Reasoning, necessity, and logic: Developmental perspectives. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Keating, D. P. & MacLean, D. J. (1987). Reconstruction in cognitive development: A post-structuralist agenda. In P. B. Baltes, D. L. Featherman, and R. M. Lerner (Eds.), Life span development and behavior (Vol. 8). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Newmann, F. M. (in press). Higher order thinking the teaching of social studies: Connections between theory and practice. In D. Perkins, J. Segal, & J. Voss (Eds.), Informal reasoning and education. Hillsdale, NJ: Lawrence Erlbaum Associates.

Oliver, D.W. & Shaver, J.P. (1966). Teaching public issues in the high school. Boston: Houghton Mifflin.

Rawls, J. (1971). A theory of justice. Cambridge, MA: Harvard University Press.

Patterson, J. H., & Smith, M. S. (1986). The role of computers in higher order thinking. In 1985 Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.

Perrone, V. & Associates. (1985). Portraits of high schools: A supplement to high school: A report on secondary education in America. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). The shopping mall high school: Winners and losers in the educational marketplace. Boston: Houghton Mifflin.

Resnick, L. B. (1987). Education and learning to think. Washington, DC: National Academy Press.

Schrag, F. (1987). Evaluating thinking in school. Contact author, National Center on Effective Secondary Schools, University of Wisconsin, Madison, WI 53706.

Schrag, F. (in press). Thinking in school and society. London: Routledge & Kegan Paul.

Stake, R.E. & Easley, J.A. (1978). Case studies in science education. Washington, DC: National Science Foundation.

Sternberg, R. J. (1983). Components of human intelligence. Cognition, 15, 1-48.

Table 1

Rater Reliability
(Pearsonian Correlations)*

School	Visit 1 Lesson		Visit 1 Pooled	Visit 2		Visit 2 Pooled	Visit 3		Visit 3 Pooled	Mean of Pooled Across Visits
Arnold	1	.77		.87			.87			
	2	.90	.88	.85		.90	.97		.91	.90
	3	.93		.91			.90			
Bradley	1	.78		.79			.90			
	2	.76	.79	.90		.84	.87		.90	.84
	3			.80			.90			
Carlsberg	1	.84		.82			.95			
	2	.87	.81	.93		.90	.91		.92	.88
	3	.70		.96						
Grandville	1	.94		.95		.93				.94
	2	.91	.94	.90						
	3	.96								
Scarborough	1	.51		.96			.70			.81
	2	.57	.74	.74		.86	.83		.82	
	3	.68		.88			.87			

Mean across lessons = .85
Mean of pooled data = .87

*Correlations of two ratings on 15 five-point scales and 2 four-point scales for each lesson. Pooled correlations are based on all ratings per school visit.

11-24

Table 2

Inter-Rater Agreement for Each Scale*

Scale	% Exact Agreement	% Differ by 1 or Less
1	59	93
2	64	100
3	46	97
4	67	100
5	54	95
6	59	92
7	54	97
8	46	90
9	85	92
10	64	97
11	56	97
12	64	95
13	62	97
14	82	100
15	87	100
16	92	100
17	64	100

* % agreement is based on 39 ratings per scale by two raters. Pairs of raters varied among 4 researchers.

Table 3

Nature of Discourse in 160 High School Social Studies Lessons (5 high schools)

Classes were rated from 1 - 5. 1 = "very inaccurate" description of class; 5 = "very accurate." Findings are reported as the percent of lessons receiving each rating.

		1	2	3	4	5
*1. FEWTOP	In this class, there was sustained examination of a few topics rather than a superficial coverage of many.	1	8	22	29	41
*2. SUBCOH	In this class, the lesson displayed substantive coherence and continuity.	1	23	28	28	20
*3. TIME	In this class, students were given an appropriate amount of time to think, that is, to prepare responses to questions.	0	2	9	44	44
4. STUEXP	In this class, students' personal experience (where relevant) was integrated into the lesson.	69	21	7	3	1
5. TCONS	In this class, the teacher carefully considered explanations and reasons for conclusions.	10	32	30	15	13
*6. TCHAL	In this class, the teacher asked challenging questions and/or structured challenging tasks (given the ability level and preparation of the students).	1	12	16	32	39
7. TSOC	In this class, the teacher pressed individual students to justify or to clarify their assertions in a Socratic manner.	47	28	15	6	5
8. TORIG	In this class, the teacher tried to get students to generate original and unconventional ideas, explanations, or solutions to problems.	40	38	14	6	3
9. TAWAR	In this class, the teacher showed an awareness that not all assertions emanating from authoritative sources are absolute or certain.	64	13	17	3	4
*10. TMOD	In this classroom, the teacher was a model of thoughfulness. (Principal indications are: the teacher showed appreciation for students' ideas and appreciation for alternative approaches or answers if based on sound reasoning; the teacher explained how he (she) thought through a problem, the teacher acknowledged the difficulty of gaining a definitive understanding of the topic.	1	21	29	28	20
11. SCRIT	In this class, students assumed the roles of questioner and critic.	14	24	25	16	21
*12. SEXPL	In this class, students offered explanations and reasons for their conclusions.	4	26	20	34	15
13. SORIG	In this class, students generated original and unconventional ideas, explanations, hypotheses or solutions to problems.	17	13	48	22	1
14. SARTI	In this class, student contributions were articulate, germane to the topic and connected to prior discussion.	1	14	14	51	20

		0-25%	25-50%	50-75%	75-100%
15. TALK	What proportion of students were active participants?	4	15	40	41

		0%	1-25%	25-50%	50-75%	>75%
16. DISC	What proportion of time did students spend engaged in thoughtful discourse with each other?	59	28	6	5	3

		25%	25-50%	50-75%	>75%
17. INVO	What proportion of students showed genuine involvement in the topics discussed? (Cues include raising hands, attentiveness manifested by facial expression and body-language, interruptions motivated by involvement, length of student responses).	14	28	29	29

---

*These variables are considered minimal requirements for a thoughtful lesson.

Table 4

Correlations Among Classroom Observation Scales

($p < .05$, N = 160)

	1 FEWTOP	2 SUBCOH	3 TIME	4 STUEXP	5 TCONS	6 TCHAL	7 TSOC	8 TORIG	9 TAWAR	10 TMOD	11 SCRIT	12 SEXPL	13 SORIG	14 SARTI	15 TALK	16 DISC	17 INVO	18 HOTAV*
1. FEWTOP	1.00	.44	-	-	.42	.47	.28	.30	-	.42	-	.45	.31	.27	-	-	.30	.73
2. SUBCOH		1.00	-	-	.62	.28	.44	-	.24	.64	-	-	.28	.24	-	-	-	.67
3. TIME			1.00	-	-	-	-	-	-	-	-	-	-	.22	-	-	-	-
4. STUEXP				1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5. TCONS					1.00	.44	.61	-	-	.79	-	.43	.25	.37	-	-	.22	.71
6. TCHAL						1.00	.34	.35	-	.50	.27	.60	.36	.42	.30	.24	.43	.74
7. TSOC							1.00	-	-	.52	-	.33	.24	.28	.21	-	-	.51
8. TORIG								1.00	-	-	.38	.22	.48	.35	.35	.37	.46	.28
9. TAWAR									1.00	.21	-	-	-	-	-	-.22	-	-
10. TMOD										1.00	-	.45	.33	.41	-	-	.32	.79
11. SCRIT											1.00	.43	.42	.60	.46	.39	.52	-
12. SEXPL												1.00	.46	.63	.43	.24	.51	.73
13. SORIG													1.00	.67	.55	-	.48	.48
14. SARTI														1.00	.52	-	.57	.56
15. TALK															1.00	.35	.51	.34
16. DISC																1.00	.28	-
17. INVO																	1.00	.44
HOTAV																		1.00

* HOTAV = sum of fewtop, subcoh, time, tchal, tmod, sexpl.

Table 5

Principal Components Analysis of 17 Variables of Thoughtfulness*

ROTATED LOADINGS

	1	2	3
SCRIT	0.812	-0.140	-0.006
SARTI	0.766	0.322	0.215
INVO	0.740	0.219	0.041
TALK	0.725	0.107	0.078
SORIG	0.697	0.291	0.256
SEXPL	0.605	0.481	0.001
TORIG	0.576	0.162	-0.343
DISC	0.530	-0.130	-0.597
TCONS	0.031	0.870	0.133
TMOD	0.118	0.848	0.234
SUBCOH	-0.085	0.777	0.183
TSOC	0.052	0.698	0.032
FEWTOP	0.238	0.632	-0.272
TCHAL	0.429	0.594	-0.172
TAWAR	-0.007	0.148	0.659
STUEXP	0.193	-0.024	0.400
TIME	0.253	-0.057	-0.056

VARIANCE EXPLAINED BY ROTATED COMPONENTS

1	2	3
4.150	3.886	1.401

PERCENT OF TOTAL VARIANCE EXPLAINED

1	2	3
24.409	22.850	8.241

* 3 components were retained and rotated through VARIMAX

Table 6

Comparisons Between Top and Bottom Quintiles of Lesson Thoughtfulness (HOTAV) on All Scales

Variable	High (N=36) Mean$_1$	(S.D.)$_1$	Low (N=37) Mean$_2$	(S.D.)$_2$	SD_p**	$\bar{x}_1-\bar{x}_2/SD_p$**	t	Prob.
HOTAV	4.60	(.21)	2.78	(.23)	.68	2.68	35.75	.000
*1. FEWTOP	4.38	(.38)	2.92	(.89)	.99	1.93	11.86	.000
*2. SUBCOH	4.53	(.70)	2.46	(.77)	1.09	1.90	12.05	.000
*3. TIME	4.56	(.56)	4.16	(.83)	.72	.56	2.36	.021
4. STUEXP	1.56	(1.08)	1.35	(.59)	.81	.26	1.01	.318
5. TCONS	4.19	(.82)	1.86	(.71)	1.18	1.97	12.94	.000
*6. TCHAL	4.78	(.42)	2.60	(.83)	1.07	2.04	14.08	.000
7. TSOC	2.83	(1.80)	1.19	(.40)	1.14	1.44	8.00	.000
8. TORIG	1.94	(.80)	1.27	(.51)	.99	.68	4.35	.000
9. TAWAR	1.67	(1.17)	1.24	(.50)	1.09	.39	2.02	.047
*10. TMOD	4.67	(.54)	2.32	(.58)	1.07	2.20	17.93	.000
11. SCRIT	2.94	(1.17)	2.70	(1.39)	1.34	.18	.80	.425
*12. SEXPL	4.22	(.59)	2.19	(.84)	1.14	1.78	11.89	.000
13. SORIG	3.08	(.97)	2.00	(1.00)	1.01	1.07	4.63	.000
14. SARTI	4.33	(.59)	3.08	(1.09)	.95	1.32	6.09	.000
15. TALK	3.42	(.65)	2.70	(1.00)	.84	.86	3.61	.001
16. DISC	1.58	(.81)	1.38	(.54)	.98	.20	1.28	.206
17. INVO	3.19	(.86)	2.05	(.97)	1.04	1.09	5.32	.000

* HOTAV is the sum of these items.

** SD_p is the standard deviation of all 160 lessons in the sample.

Table 7

Frequencies of Classroom Practices
High and Low Lessons

Variable	High (N=36)	Low (N=42)	Pearson Chi-Square	Probability
DLECT	1	18	19.94	.000
DFILM	0	0	---	---
DRECI	0	14	16.85	.000
DTCHD	33	16	19.39	.000
DSTUD	0	1	.99	.321
DGRUP	8	2	4.62	.037
DSEAT	2	3	.19	.666
DREPT	0	0	---	---
DOTHR	0	0	---	---
RTEXT	8	20	7.82	.005
RARTC	2	4	.67	.414
RPRIM	13	5	5.02	.025
RLITR	0	0	---	---
ROTHR	12	6	2.88	.090
RNONE	6	9	.66	.418
WOUTL	0	0	---	---
WNOWN	24	19	1.77	.184
WWORK	11	12	.03	.863
WCOMP	7	3	1.98	.159
WPARA	1	3	1.00	.317
WPAPR	0	0	---	---
WOTHR	1	0	1.04	.307
WNONE	4	7	.87	.351

Table 8

School Differences in HOTAV

	SCHOOL									
	Arnold		Bradley		Carlsberg		Grandville		Scarborough	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
HOTAV	3.91	.60	3.50	.68	4.13	.48	3.99	.57	3.15	.58

Table 9

Regression of HOTAV on Schools*
(N=160)

MULTIPLE R: .528 SQUARED MULTIPLE R: .278
ADJUSTED SQUARED MULTIPLE R: .260 STANDARD ERROR OF ESTIMATE: 0.586

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	T	P(2 TAIL)
CONSTANT	3.151	0.105	0.000	29.910	0.000
ARA	0.764	0.147	0.455	5.206	0.000
BRK	0.344	0.148	0.203	2.329	0.021
CAR	0.980	0.148	0.577	6.629	0.000
GRN	0.839	0.148	0.494	5.677	0.000

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	20.562	4	5.140	14.945	0.000
RESIDUAL	53.313	155	0.344		

*Lessons were assigned dummy values, 1 or 0, for each of the variables: ARA (Arnold), BRK (Bradley), CAR (Carlsberg), GRN (Grandville). Scarborough was the "constant." Coefficients represent the difference in the HOTAV mean between each school and Scarborough; for example, Arnold exceeds Scarborough's mean by .76.

Table 10

Correlations of HOTAV with
Lesson Background Variables

(N=160)

% of Students	
9th	- .08
10th	.24
11th	- .22
12th	.15
low 1/3	- .07
mid 1/3	.07
high 1/3	- .01
black	- .36
hispanic	.40
Asian	.18
other minority	- .05
course required (1) or elective (0)	.02

Table 11

Distributions of the Percent of Minority Students in 160 Lessons

	BLACK	ASIAN	HISPANIC	OTHER
Minimum	0	0	0	0
Maximum	81.0	42.0	15.0	8.0
Mean	16.6	10.0	2.4	0.3
Standard Deviation	24.2	10.9	3.5	1.3

Table 12

Regression of HOTAV on Lesson Background Variables

MULTIPLE R: .370 SQUARED MULTIPLE R: .137
ADJUSTED SQUARED MULTIPLE R: .114 STANDARD ERROR OF ESTIMATE: 0.641

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2TAIL)
CONSTANT	4.071	0.574	0.000	1.0000000	7.090	0.000
GRADEAV	0.002	0.053	0.003	.9666467	0.043	0.966
ABILAV	-0.085	0.091	-0.072	.9384370	-0.932	0.353
BLACK	-0.011	0.002	-0.381	.9336964	-4.930	0.000
DUMREQ	0.013	0.123	0.008	.9675763	0.108	0.914

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	10.101	4	2.525	6.138	0.000
RESIDUAL	63.774	155	0.411		

Table 13

Regression of HOTAV on Schools* and Lesson Background Variables
(N=160)

MULTIPLE R: .558 SQUARED MULTIPLE R: .311
ADJUSTED SQUARED MULTIPLE R: .275 STANDARD ERROR OF ESTIMATE: 0.580

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	T	P(2 TAIL)
CONSTANT	4.450	0.673	0.000	6.609	0.000
ARA	0.552	0.296	0.329	1.866	0.064
BRK	0.153	0.287	0.090	0.535	0.594
CAR	0.960	0.245	0.565	3.919	0.000
GRN	0.649	0.328	0.382	1.982	0.049
GRADEAV	-0.083	0.053	-0.119	-1.582	0.116
ABILAV	-0.050	0.095	-0.042	-0.525	0.601
BLACK	-0.004	0.005	-0.150	-0.877	0.382
DUMREQ	-0.171	0.123	-0.105	-1.392	0.166

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	22.993	8	2.874	8.530	0.000
RESIDUAL	50.882	151	0.337		

*Lessons were assigned dummy values, 1 or 0, for each of the variables: ARA (Arnold), BRK (Bradley), CAR (Carlsberg), GRN (Grandville). Scarborough was the "constant." Coefficients represent the difference in the HOTAV mean between each school and Scarborough; for example, Arnold exceeds Scarborough's mean by .55.

II - Appendix

Observational Scales for Assessing Higher Order Thinking in High School Social Studies Classes

For scales 1-14, classes were rated from 1 - 5. 1 = "very inaccurate description of this class; 5 = "very accurate" description of this class. Technical notes on the scoring of the scales will be made available at the conclusion of the study.

GENERAL

1. There was sustained examination of a few topics rather than superficial coverage of many.

Mastery of higher order challenges requires in-depth study and sustained concentration on a limited number of topics or questions. Lessons that cover a large number of topics give students only a vague familiarity or awareness and, thereby, reduce the possibilities for thoughtful study.

2. The lesson displayed substantive coherence and continuity.

Intelligent progress on higher order challenges demands systematic inquiry that builds on relevant and accurate substantive knowledge in the field and that works toward the logical development and integration of ideas. In contrast, lessons that teach material as unrelated fragments of knowledge, without pulling them together, undermine such inquiry.

3. Students were given an appropriate amount of time to think, that is, to prepare responses to questions.

Thinking takes time, but often recitation, discussion, and written assignments pressure students to make responses before they have had enough time to reflect on their responses.
Promoting thoughtfulness, therefore, requires periods of silence where students can ponder the validity of possible responses.

TEACHER BEHAVIOR

4. The teacher asked challenging questions and/or structured challenging tasks (given the ability level and preparation of the students).

By our definition higher order thinking occurs only when students are faced with questions or tasks that demand analysis, interpretation, or manipulation of information; that is, non-routine mental work. In short, students must be faced with the challenge of how to use prior knowledge to gain new knowledge, rather than the task of merely retrieving prior knowledge.

5. The teacher carefully considered explanations and reasons for conclusions.

The resolution of higher order challenges often depends not simply upon offering explanations and reasons, but upon the quality of explanations or reasoning given to support conclusions. Rigorous analysis and evaluation of reasoning, therefore, is central to the promotion of thoughtfulness.

6. The teacher pressed individual students to justify or to clarify their assertions in a Socratic manner.

Socratic interchanges probe an individual's level of understanding. Such dialogue presses the student to consider the validity of evidence and reasoning, to suggest alternative perspectives that may otherwise escape attention, and to identify rather rigorously what one does and does not know. These probes usually constitute their own higher order challenges as they facilitate development of knowledge on the topic in question.

7. The teacher encouraged students to generate original and unconventional ideas, explanations, or solutions to problems.

Higher order challenges can require creative, intuitive insights and alternative perspectives that diverge from conventional knowledge or from expected lines of reasoning. Promoting this sort of mental flexibility will help students to cope not only with ill-structured problems that may explicitly invite creativity, but also with well-structured ones that may require unconventional thinking.

8. The teacher showed an awareness that not all assertions emanating from authoritative sources are absolute or certain.

Curiosity and skepticism about the creation of knowledge and the nature of truth are dispositions of thoughtfulness that help to sustain higher order thinking. For these dispositions to be promoted, authoritative sources must not always be accepted uncritically. Instead, students must be made aware that what may pass as final, conclusive knowledge may often be problematic and subject to future revision and development.

9. Students' personal experience (where relevant) was integrated into the lesson.

Student engagement with and comprehension of a problem can often be enhanced if inquiry on the problem is related to students' actual life experiences.

10. The teacher was a model of thoughtfulness.

To help students succeed with higher order challenges, teachers themselves must model thoughtfulness as they teach. Of course, a thoughtful teacher would demonstrate many of the behaviors described

above, but this scale is intended to capture a cluster of additional characteristics likely to be found in any thoughtful person. Key indicators include showing interest in students' ideas and in alternative approaches to problems; showing how he/she thought through a problem (rather than only the final answer); and acknowledging the difficulty of gaining a definitive understanding of problematic topics.

STUDENT BEHAVIOR

11. Students offered explanations and reasons for their conclusions.

The answers or solutions to higher order challenges are rarely self-evident. Their validity often rests on the quality of explanation or reasons given to support them. Therefore, beyond offering answers, students must also be able to produce explanations and reasons to support their conclusions.

12. Students generated original and unconventional ideas explanations, hypotheses or solutions to problems.

This is a counterpart to scale #8. Higher order thinking is more evident in classrooms in which teachers not only encourage original responses from students, but where students also supply them.

13. Students assumed the roles of questioner and critic.

As students work toward answers for higher order challenges, they are more likely to succeed if they have opportunities to ask questions about and to criticize proposed approaches and answers. In contrast, if students act primarily as passive recipients of information, they are not likely to develop the level of understanding needed to carry out useful analyses and interpretations.

14. Student contributions were articulate, germane to the topic and connected to prior discussion.

In a formal sense, students may respond actively with assertions, explanations, questions, and creative ideas, but these may be inarticulate and irrelevant to the topic being examined. Just as scale #2 requires substantive coherence and continuity in the lesson as a whole, student participation cannot be considered thoughtful unless it is reasonably articulate and relevant.

15. What proportion of students participated verbally in the lesson?

To solve higher order challenges, students must actively use their minds. This cannot be observed directly, but one indicator is the extent to which they speak in class about the subject being examined. That this, the more students who produce discourse on a topic the greater the likelihood of higher order thinking by the class as a whole.

16. What proportion of time did students spend engaged in thoughtful discourse with one another?

Because of the importance of dialogue in promoting higher order thinking, yet the logistical difficulty of the teacher responding to the individual ideas of each student, it can be useful for students to talk with one another to test and refine their ideas. The amount of thoughtful discourse among students can, therefore, be seen as one indicator of thoughtful discourse in the class as a whole.

17. What proportion of students showed genuine involvement in the topics discussed?

To make progress on higher order challenges, students must give their attention, concentration, and mental effort. We summarize these dimensions of involvement as student engagement. One indicator is observable verbal participation, but it is also possible to become intensely involved while listening. This scale assesses the extent of student engagement based on cues that extend beyond speaking. Principal indicators include raising hands, attentiveness manifested by facial expression and body-language, interruptions in the discussion to raise a substantive point or question, the length of student responses.

III

EXPLORING ASPECTS OF TEACHERS' THINKING ABOUT PROMOTING STUDENTS' THINKING

Joe Onosko

INTRODUCTION:

A prominent theme of educational reform in the 1980's is to develop students' "higher order" and "critical" thinking. This has been reflected in numerous: a) national reports advocating increased emphasis on thinking in schools (Adler, 1982; Boyer, 1983; College Board, 1983; Sizer, 1984; etc.); b) scholarly works that attempt to conceptualize thinking; c) curriculum materials developed to help teachers promote students' thinking; d) workshops and programs on thinking sponsored by school districts and professional organizations throughout the country; and, e) national and state efforts to incorporate test items requiring higher level thinking on achievement tests and other standardized assessment instruments.

What is strikingly absent in the literature and in programmatic reform efforts is an attempt to understand and learn from the work of practitioners. This may be due in part to the bleak portrait of actual classroom practice painted in the literature (Cuban, 1984; Goodlad, 1984; Hoether and Ahlbrand, 1969; Perrone, 1985; Sirotnik, 1983). Nonetheless, we know that some teachers do challenge students intellectually (Lightfoot, 1983; Sizer, 1984; Swartz, 1987). If we are to understand how to promote thinking in the classroom, we need to learn more about how teachers think about their work. Over the past ten years research has documented that teachers' thoughts and beliefs play a crucial mediating role in their classroom practice (Clark & Peterson, 1987; Peterson, Fennema, & Carpenter, 1987; Shavelson & Stern, 1981; Schulman, 1987; Zeichner & Teitelbaum, 1982). As an exploratory effort we compare the thought of teachers who quite consistently promote thinking in their classrooms with the thought of teachers who do so less consistently. Should important differences emerge in the thinking of the two groups, these differences may provide clues to influencing classroom practice through changing teachers' thinking about practice. However, if only similarities are found in the thinking of the two groups then no clues can be offered.

A sample of 10 high school teachers were selected from a total pool of 15 teachers from 5 secondary schools (3 teachers per school). Each teacher was observed on 9 occasions.[1] Three class observations took place during each of the 3, 4-day school-site visits spread over the course of the 1986-87 academic year (i.e., fall, winter, spring). In addition, at least 2 hours during each of the three site-visits was spent in interview with each teacher to probe their thinking about teaching, students, colleagues, etc. The teachers are members of social studies departments that have made conscious efforts and progress at fostering students'

thinking. The 5 schools were selected through a nation-wide search in which 60 promising social studies departments were nominated. Through extended phone conversations with the principal, department head, and one or two teachers, and a one-day site visit, judgments were made regarding the degree to which each department fulfilled criteria specified in the nomination request announcement. Criteria included a department-wide emphasis on thinking as witnessed in the practices of department members and in conversations with them, and that thinking was emphasized in required courses and in courses with a preponderance of low-achieving, low-SES, or minority students. The teachers in the pool of 15 are considered by the department chairs to be among the best in their respective departments at emphasizing thinking with students.[1]

The 5 teachers with the highest percentage of lessons receiving a score of "4" or "5" on each of the 6 observational scales representing the minimal requirements of a thoughtful discourse lesson were designated "high scorers."[2] See Newmann (this volume) which provides a rationale for and description of the observational scales used to measure the thoughtfulness of classroom discourse. The 5 teachers with the lowest percentage of lessons receiving a moderately high score of 3 or more on each of the 6 observational scales were designated "lower scorers."[3] The remaining or "middle" 5 teachers in the pool of 15 were excluded from the analysis.

---

[1]Five of the 15 teachers were observed on eight rather than nine occasions due to personal illness, a snow day, etc.

[2]Because all 6 scales are viewed as necessary components for thoughtful discourse, a high score (i.e., "4" or "5") was required on each scale before a teacher's lesson was to be considered outstanding. Basing teacher performance on the mean score of a lesson across all six scales was not used. The reason for this is that a low score of "2" or "3" on a few scales combined with 5's on the remaining scales could result in a higher mean score across the six scales than a lesson receiving all 4's--leading to the possible exclusion of the latter as an exemplary lesson.

[3]Due to the skewedness of the sample, the lower scoring group is considered less-outstanding rather than non-outstanding with respect to the high scorers. Therefore selection of lower scorers was based upon the lowest percentage of good lessons (i.e., a score of "3" or more on each of the 6 scales) rather than the highest percentage of poor lessons. It should be noted that the third observational scale (i.e., teacher gave students time to think) required a score of "4" or more rather than "3" or more, as otherwise all but 18 of the 135 lessons would qualify as moderately good lessons on this dimension of teachers' practice.

Teachers were unaware of the specifics of the study's research agenda. They did not know that two groups of teachers would be compared on various dimensions of their thinking about teaching, students, colleagues, etc. Nor were teachers aware of the observational scales used to measure dimensions of thoughtfulness in their lessons. Researchers who summarized teacher responses did not know which teachers from the pool of 15 would comprise the two groups of high and lower scorers until statistical analyses were performed months after the data was gathered.

Five areas of teachers' beliefs and theories were explored: 1) teachers' conceptions of thinking; 2) the relative importance given to thinking as an instructional goal; 3) their views on the issue of depth vs breadth of content coverage; 4) their perceptions of organizational barriers to promoting thinking in classrooms; and, 5) their views of students. Our selection of these 5 areas was based in part on previous research.

Previous research indicates, for example, that teachers possess a variety of theories and beliefs about instruction (e.g., ideas on learning, reading, classroom management, student motivation, etc.), and that these theories and beliefs help guide their instructional efforts (Clark & Peterson, 1987; Marland, 1977; Munby, 1983). Indeed, it is difficult to imagine a teacher pursuing an instructional goal without some image, set of ideas, or conception of the goal being pursued. We were therefore curious to see if high scorers: a) possessed different and/or more elaborate conceptions of thinking than lower scorers; and, b) placed greater emphasis on thinking as an instructional goal than lower scorers.

Turning to the issue of depth vs breadth of content coverage, at one end of the spectrum are teachers who rigorously pursue breadth of coverage, galloping their learning hard over numerous content pastures at a mind-numbing pace and reducing students' educational experience to simplistic memorization of names, dates, places, events, etc. At the other end of the spectrum are teachers who rigorously pursue depth of coverage, chaining their learning hard to a few content pastures and reducing students' educational experience to detailed understanding of a relatively narrow range of human experience. Is either approach more likely to promote the development of students' thinking? The conceptual literature on thinking highlights the importance of thinking skills, dispositions, and content understanding for effective thinking performance (Newmann, 1988a; Siegel, 1988). Presumably, time spent developing students' thinking skills and dispositions, and their understanding of the content under study, will reduce time for content coverage. Some educators have in fact argued that there must be a movement away from massive coverage if students are to develop an ability to thinking deeply and effectively about topics and issues (Newmann, 1988b; Sizer, 1984). How do teachers attempting to promote students' thinking tackle this dilemma? We were curious to see if high and lower scorers addressed or

attempted to resolve the dilemma of depth vs breadth of coverage differently.

The promotion of thinking also seems related to organizational features of teachers' work settings. For example, large class size, large total student loads, relatively short 45 minute class periods may severely constrain teachers' instructional efforts (Powell et al., 1986; Sizer, 1984). The adulteration of academic tasks through the process of "negotiation" between teacher and students (McNeil, 1983) may, in part, be due to these debilitating organizational features. Indeed, one respected observer of the classroom has argued that the development of students' higher order thinking and organization features of the traditional classroom are incompatible if not mutually exclusive (Cuban, 1984b). Some research suggests that frustration over working conditions is a primary reason teachers leave the profession (Goodlad, 1984). We were curious to see if high and lower scorers perceived or attempted to modify various organizational barriers differently. More specifically, do high scorers possess more effective strategies for dealing with these barriers?

Finally, we looked at teachers' perceptions of students. Much research has documented the powerful impact teachers' perceptions of students have on the instructional process, including the number and kind of classroom interactions between teacher and student (King, 1980; Page, 1984; Peterson & Barger, 1984), the types of instructional goals pursued and the way classroom behavior is managed (Brophy & Rohrkempter, 1981), and teachers' expectations of student performance (Good, 1987). We wanted to see if high scorers' perceptions of students are in any ways different from the perceptions of students held by lower scorers.

COMPARISON OF HIGH AND LOWER SCORERS ON BACKGROUND CHARACTERISTICS AND ORGANIZATION DIMENSIONS OF THEIR WORK:

Before presenting the findings, a brief comparison of the two groups is provided on selected background characteristics and organizational dimensions (see Table 1). Such a comparison may respond to suggestions from previous research that factors of class size, ethnic background, prior achievement, etc., affect learning outcomes and instructional practices. Do high scorers in our sample have a more favorable teaching situation with respect to the above variables? If major differences between the two groups fail to emerge on these variables then explanations for observed differences in the instructional practices of the two groups will need to rely on other sources.

Four of five teachers in both the high and lower scoring groups are male. This percentage closely parallels national statistics in which women constitute only about 26% of high school social studies teachers (Rutter, 1985). The high scorers averaged 14.6 years of social studies teaching experience, while the mean

for the lower scorers was 17.0. However, experience varied greatly within each group; 2 to 21 years for the high and 3 to 25 years for the lower scorers. The amount of formal education is comparable for the two groups of teachers. Most of the high scorers (4) have a Masters' degree plus additional credits, while the remaining member has a Masters' degree. Among the lower scorers, 1 has a Doctorate, 2 possess a Masters' plus additional credits, and 2 earned a Masters' degree. Job satisfaction is also comparable, as most teachers (4) in both groups state they are satisfied with their job "most" of the time while the remaining member from each group derives satisfaction only "half" of the time.

Three of 5 schools are represented by teachers in the high scoring group and 4 of 5 schools in the lower scoring group, and 2 schools have teachers who appeared in each group. With respect to subject matter, 3 high scorers taught a history class while the other 2 taught social science courses (i.e., economics and psychology). Within the lower scoring group, 4 taught history and 1 a social studies course (i.e., American government, or "Civics"). A sizable difference between the two groups emerges with respect to the number of teaching periods per day. High scorers averaged 4.6 teaching periods per day whereas the lower scorers averaged one less teaching period, or 3.6. Class size however was about the same, 28.6 for the high and 26.4 students for the lower scoring group. Because of the additional class assignment high scorers' total student load was substantially higher; 131.8 students compared to only 97.6 for the lower scorers. The number of course preparations was essentially the same, 2.2 for high scorers and 2.4 for lower scorers.

Turning to student class compositions, most of high scorers (4) worked with students of diverse or low achievement levels whereas a majority of the lower scorers (3) worked with high achieving students.[4] Both groups had similar Hispanic compositions; 11% for the high and 12.8% for the lower scoring teachers. Higher scorers had more Asian students than lower scorers, 5% to 1.2%, but fewer blacks, 12.6% to 28.8%. However, the disparity between high and lower scorers regarding percentage of black students should not be overemphasized as variance was

---

[4]At each school we requested for observation three types of classes: a) primarily low achieving students; b) a diverse range of student achievers; and, c) the class exhibiting the most higher-order thinking (which usually comprised high achievers). Prior social studies achievement levels of students were determined by the department head and later confirmed by the teachers themselves. Student achievement data on teachers' classes pertains only to the particular class researchers observed for the observation portion of the study design. Data was not collected on prior achievement, ethnic, or racial compositions of teachers' total student load.

large within each group; i.e., from 0% to 37% for high and 0% to 74% for lower scoring teachers. In addition, most lower scoring teachers (3) had less than 3% black students in their respective classes.

To summarize from the above data, similarities exceed differences between the two groups of teachers. And where substantial differences do exist, lower scorers tend to have the more favorable situation (e.g., fewer # of students, fewer teaching periods, more high achieving students, etc.). This observation, however, should not be taken to imply that assigning teachers a larger student load, more low achievers, or more class periods per day will increase the level of classroom thoughtfulness. This finding may simply be an artifact of the small sample size. But if this represents a real difference, it may be interpreted to mean that teachers' instructional approaches can override otherwise constraining organizational dimensions of their work. Because this finding runs counter to all of the recommendations for teacher professional development and improvement, future research should keep this interesting anomaly in mind.

INSTRUCTIONAL GOALS FOR SOCIAL STUDIES:

A number of questions were used to identify teachers' instructional goals. The findings from any single question may not be compelling when taken in isolation, but as a group the questions generate findings that high scorers place greater emphasis on thinking than their lower scoring counterparts.

Though teachers from both the high and lower scoring groups all state that helping students become better thinkers is very important for their feelings of success, when asked to identify their highest priority goal for social studies instruction and the goals that focus their lesson planning differences between the two groups begin to emerge.[5] With respect to their highest priority goal, "critical thinking and problem solving" was given greater emphasis by high scorers; that is, most high scoring teachers (4) vs a few lower scorers (2) cited thinking and problem solving as their highest goal (see Table 2). That thinking and problem-solving was less emphasized by lower scorers could theoretically prove to be an insignificant finding, as thinking can play a prominent role in the pursuit of other goals (e.g., "teaching facts, concepts, and theories"; "developing discussion skills"; etc.). However, the above difference between the two groups gains in importance when one also considers that all high scorers (5) but only some of the lower scorers (3) cited "thinking and problem solving" as one of the many goals that focus their

---

[5]Teachers were asked to select their highest goal and the goals that focus their planning from a list of potential goals (see Table 2).

lesson planning efforts (see Table 3). Note that teachers here were encouraged to identify any and all goals that focus their lesson planning, yet one finds a few lower scorers not mentioning thinking as a planning goal.

In addition to the above findings, inconsistencies emerge in the responses of lower but not high scoring teachers. A few of the lower scorers (2) failed to list their stated highest priority goal as one of the goals that focus their lesson planning. No such omission occurred among high scorers. For example, one lower scorer identified teaching students about "past and present problems and issues faced by the U.S. and World" as her highest goal, yet the goals focusing her lesson planning were to "teach students constructive social values and foster responsible citizenship", and "teach students facts, concepts, and theories central to understanding history and the social sciences."

Inconsistencies also arose among the same two lower scorers with respect to their highest priority goal; that is, their highest priority goal changed from visit 1 to visit 3. For example, one lower scorer identified "teaching students critical thinking and problem solving" during the fall (visit 1) as her highest goal yet in the spring (visit 3) replaced it with "teaching students constructive social values and responsible citizenship." There is no indication or reason to believe that the instructional approach of either teacher had changed during the year.

In another query of teachers' goals we asked high and lower scorers if exposing students to subject matter content is, in general, "more", "equally", or "less" interesting to them as developing students' thought and reasoning processes. All of the high scorers find exposing students to subject matter content less interesting compared to only 1 among lower scorers. Most of the lower scorers (3) find content exposure equally interesting while one finds it, in fact, more interesting.[6] These findings (see Table 4) are consistent with findings cited above in which high scorers place greater emphasis on thinking as a highest priority goal and as a goal focusing their lesson planning.

Additional support for the differing emphases placed on thinking by the two groups can be culled from teachers' responses to the following general probe: "What in particular gives you satisfaction as a teacher?" Teachers' responses can be grouped into the following four kinds or categories of satisfaction: a) seeing students thinking; b) seeing students responding; c) working

---

[6]However, when teachers are asked if exposing students to content vs developing students' thinking is more or less "difficult" (rather than "interesting"), both groups (4 of 5) agreed that is was "more" difficult to develop students' thinking.

with colleagues; and, d) constructing lesson plans (see Table 5). With respect to the latter two categories (i.e., working with colleagues and constructing lessons), the responses of high and lower scorers are very similar. One teacher from each group made reference to lesson planning as satisfying, while 3 high and 2 lower scorers cited their work with colleagues as satisfying.

However, important differences appeared between the two groups regarding the other satisfaction categories (i.e., students thinking and students responding). All of the high scorers (5) cite as satisfying student behaviors one would associate with thinking (e.g., "seeing students start to make connections", "seeing students gain a more precise understanding", "students wrestling with values and making links", "leading students to grasp concepts/skills", "teaching students to generalize from data", etc.), whereas only 1 lower scoring teacher cited a satisfaction one might associate with thinking--and that rather marginally (i.e., "touching a young mind"). All of the lower scorers (5), however, say they derive satisfaction when students are responding in class (e.g., "when students show interest", "when students are generally responsive", "the responses of students", etc.), whereas only 1 high scorer made a comment indicating satisfaction when students are responding (i.e., "students commenting").

The rather profound differences between the two groups regarding student behaviors they find satisfying may simply be a matter of semantics; that is, lower scorers may actually mean thinking-type student behaviors when they say they derive satisfaction from responding-type behaviors. However, even assuming this were the case (which is not at all clear), responses indicate that high scorers identify with greater specificity the kinds of student behavior they find exciting and satisfying, and that these behaviors are typically associated with acts of thinking.

Similar findings emerged between the two groups when teachers were asked on an open-ended question to identify the kinds of thinking tasks less receptive students are likely to resist. Though responses were quite varied--both within and between groups--the tasks that high scorers mention are more readily identifiable as involving higher order thinking, again suggesting that high scorers place greater emphasis on thinking as an instructional goal. The following thinking tasks are cited by high scorers as most frequently resisted: giving reasons for statements, giving opinions, doing a critical analysis, defending one's viewpoint, engaging in value reflection, supplying metaphors, using precise language, dealing with abstractions, essay writing, and work requiring extended effort. Lower scorers on the other hand cite activities that for the most part may or may not involve thinking: doing homework, participating in discussion, reading in class, doing vocabulary assignments, anything beyond rote memory, anything without step by step instructions, essay writing, work

that is too difficult, long readings, and dealing with abstractions.

The above differences in goal orientation between high and lower scoring teachers are further revealed during an interview session and on a questionnaire item when we asked teachers to discuss their instructional goal(s). High scorers' responses are longer, more elaborate, and place greater emphasis on the development of students' thinking.

In addition, one senses that high scorers see their work directed toward a broader, more far-ranging mission that lower scorers. Besides producing improvements in students' knowledge structures and thinking abilities, high scorers hope that students will emerge from their classes transformed in more far-reaching ways. For instance, consider the way one high scorer, Hugh,[7] describes his mission:

> Content is a vehicle to teaching critical thinking, though there are certain things you want them to know when they leave economics. No matter what subject, you can get students to think about big issues...I'd like kids to always be questioning, to always be probing. You should always be on the edge, never comfortable, no matter how well you've digested the material...don't take anything at face value...I never saw a lesson where I went to school where they didn't know the 'answer' at the end of the trail...I want kids to be able to say, 'Hey, I'm a person, I can think! That's what I'm all about. I'm suppose to think.' To think is a helleva lot better than to know. To 'know' can be a dampening experience sometimes because you're not going to probe anymore, you've expended yourself.[8]

Another high scorer, Hilary, also highlights critical thinking, placing special emphasis on perspective-taking as an essential

---

[7]All high scorers' names will begin with the letter "h" as in "high", while lower scorers' names will begin with "l" as in "lower".

[8]Teachers' responses are taken from questionnaires and follow-up interviews. Some quotes are direct, while others are quoted from researcher's notes. Researchers' notes capture the ideas expressed but may not represent the teachers' comments verbatim. Wherever possible efforts were made to return to the audio-tapes to check for response accuracy.

aspects of critical thinking about social and personal values. She states:

> Until you can begin to at least temporarily put yourself in the other guy's shoes, even if you aren't going to end up there, you cannot evolve a set of social values that are good for you or for society...It's important for students to be able to step into the perspective of another as it enables them to better understand the total situation, and to defend their own position if they still maintain it after perspective taking...The roots of prejudice and discrimination may lie in an inability to see other perspectives...in an over-programmed, media- centered world, decisions and solutions to problems are too often 'made' for people. Students develop a pattern of letting someone else do their thinking. There is a need to teach thinking so that kids will think for themselves...I try to get students to question and formulate positions...to be able to explain and support ideas...to see that question asking is a sign of intelligence not ignorance...to get them to know that they don't know and want to find out.

Except for one member, lower scorers' responses to the question of instructional goals were generally much shorter and lacked the impassioned elaboration of high scorers--regardless of the goal cited. Some lower scorers simply parroted from the list of goals we offered as examples on the questionnaire. Low scorers' responses include the following:

> Critical thinking and problem solving. Developing constructive social values. I have become far more subject oriented. I started teaching in the late 60's and 70's. I find there is a far greater need to provide students with a sound data and concept base. [Lisa].

> Teaching facts, concepts, and theories of the social sciences. Our district is emphasizing the development of discussion skills, public speaking, and things of that nature. We emphasize discussion within a convergent framework in social studies...I take my directions more or less from the district...they tell us what it is they expect from us...now they are moving toward higher level thinking in the curriculum and getting beyond the literal level...In my curriculum

> right now it's not in there but in some curriculums it is being implemented...I think it is their responsibility as an administrator to make known to me what is being taught or should be taught. [Lloyd].

To summarize, high scorers are more likely to identify thinking as their highest priority goal and as a goal focusing their planning. Also, high scorers, when asked what gives them satisfaction as a teacher, are likely to provide more detailed descriptions than lower scorers of the kinds of student behavior they find satisfying, and that these descriptions are more likely to be associated with acts of thinking. Unlike most of the lower scorers who find exposing students to subject matter content equally or more interesting than developing students' thinking and reasoning, high scorers unanimously prefer to develop students' thinking. In addition, contradictions emerged in the responses of lower scoring teachers but not high scorers; that is, the stated highest goal of two lower scorers changed over a 6 month time period without explanation, and these same two failed to mention their highest priority goal among the goals that focus their lesson planning. Finally, the goal statements of high scorers are much lengthier and more detailed, focus more on thinking, and suggest that high scorers desire to affect students in far-ranging, character-like ways. When combined, the findings from the above research probes form a montage that illustrates a more consistent instructional emphasis on thinking among high scorers compared to lower scorers.

BARRIERS TO THE PROMOTION OF THINKING:

As in the goals section above, a variety of questions were used to describe teachers' perceptions of instructional and organizational barriers to promoting students' thinking. Again, the sum of the parts paint a clearer portrait than does any single part. In this portrait, however, similarities outweigh differences between the two groups of teachers.

Teachers were asked to indicate on a scale of 1 to 5 (extremely negative to extremely positive) the effect various organizational barriers have on their classroom efforts to promote students' thinking (see Table 6). Both groups indicated that "large class size" and "large student load" had the most negative impact on their instructional efforts, with the mean score for both groups well below 2.0. The only other organizational feature that was viewed by both groups as having a negative effect on the promotion of students' thinking is "students having to take too many courses." The only organizational feature to receive a positive rating (i.e., above 3.0) by both groups of teachers was "4-5 class sessions per week". The two groups displayed a sizeable difference in their ratings only with respect to their attitudes about "ability grouping". High scorers find ability grouping

somewhat detrimental to their instructional efforts (x=2.8), while lower scorers find ability grouping somewhat beneficial (x=3.8).

Similar findings emerge when teachers were asked to identify the 3 barriers that most inhibit their attempts to promote students thinking (see Table 7). Many high scorers (3) and most lower scorers (4) cited "large student load" as one of the three most inhibiting barriers. Both groups (3 teachers each) also cited "large class size." The data suggests that teachers are not frustrated by the number of classes they must teach, as only 1 teacher in the entire sample of 10 identified "number of classes" as detrimental to their instructional efforts. Rather teacher concern focuses on the large number of students they must address in each class and over the course of the day. Aside from the issue of student load, teachers' responses reveal great diversity in the barriers they find most inhibiting.[9]

The most important finding to emerge from teachers' thoughts on barriers may not be what they identify, but rather how little they have been able to reduce the impact of these barriers. Once teachers identified the three most detrimental barriers to their classroom efforts, we asked them what, if anything, they had tried to do to reduce the effect of these barriers and with what results. Of the 15 barriers identified by each group (i.e., 3 from each teacher), high scorers' responses indicate that no meaningful change or improvement had taken place with respect to 12 of the barriers, while lower scorers' responses indicate no meaningful change with respect to 11 of 15 barriers. Some teachers shared with us their feeble attempts to reduce the impact of these barriers. For example, with respect to "large class size" one teacher makes "a special effort to incorporate non-volunteers" by calling on them during discussion, or with respect to the barrier of "large total student load" another teacher gives an occasional oral rather than written quiz to cut down on the amount of paper correction time.

The few instances where teachers did achieve some positive effect on barriers are not dramatic. For example, one high scorer was able to reduce his student load by securing a curriculum development position for one period a day, effectively reducing his class assignment load from 5 to 4. However, this is only a one year appointment and only open to one teacher in the district per year. At another school a lower scorer has reduced his student load during class discussions by occasionally having a colleague take half of the students to another classroom for discussion. Of course, the net effect is to increase the student load of the

---

[9]Pressure to cover course content emerges as an important concern of high scorers. This barrier, however, will receive separate attention in the section on depth vs breadth of content coverage.

obliging colleague. In only one situation was a teacher able to significantly reduce student load, and this by dropping to half-time employment which gave her half a load. Needless to say, this is not an option available to most teachers.

One cannot fail to notice that all of the teachers' efforts to reduce the impact of organizational barriers were attempted within the parameters defined by the present school structure; that is, teachers did not attempt to modify structural aspects the school institution. We did not formally collect data on the reasons for this situation, but based upon teachers' responses to a related question it seems that teachers are extremely resigned and pessimistic about the likelihood of positive change in the organizational features of schools. [10]

Teachers were also given the opportunity to imagine a more ideal teaching situation. They were asked to identify 3 "wishes" they would make to aid them in their efforts to promote students' thinking. Again, the concern over student load was dominant as 4 of 5 teachers in both the high and lower scoring groups requested fewer students as one of their three wishes. Aside from the problem of student load, the high and lower scoring groups have different wishes. Little consensus occurred within groups either (see Table 8).

Turning from teachers' perceptions of various organizational barriers to their perceptions of department colleagues, both high and lower scorers unanimously agree that their colleagues support rather than undermine their efforts to emphasize thinking with students. With respect to administrative authorities (i.e., department head, principal, superintendent, etc.,), most of the high (4) and lower (4) scorers feel free to diverge from the conception and set of practices the above authorities espouse for promoting students' thinking. This suggests that teachers do not perceive administrators as undermining their instructional efforts.

As for testing as a potential barrier, all high scorers (5) and most lower scorers (4) agree that departmental, school-wide, and district tests have "no effect" on their efforts to promote students' thinking in the classroom. Many of the "no effect" responses are due to the fact that 3 of 5 schools in the sample do

---

[10]For example, teachers were asked to identify 3 "wishes" they would make to aid them in their efforts to promote students' thinking. They were then asked to explain what would have to happen in their school for the wish to be fulfilled over the next 5 to 10 years. Typical teacher responses were: "...the likelihood of this is nil", "...we have had no reduction in 20 years", "...nothing is going to happen", "...little will change", "...chances: 1 in a million", etc. For details of teachers' wishes see Table 8 below.

not administer departmental, district, or school-wide tests. Differences between high and lower scorers did emerge, however, when teachers were asked to evaluate the effects of state and Advanced Placement exams on their instructional efforts. Of the 5 teachers having had experience with AP courses (i.e., 4 lower scorers and 1 high scorer), the 4 lower scorers feel that AP tests "support" a thinking emphasis while the one high scorer believes such tests "inhibit" his classroom efforts to promote students' thinking. Likewise, at the one school administering a state test, the two high scorers think the test inhibits a thinking emphasis while their lower scoring colleague sees the same test as supportive. As will be explained later in this chapter, differences here between high and lower scorers may be attributed to differences in their instructional goals and their resolution of the depth vs breadth dilemma.

To summarize findings on organizational barriers, the primary concern of teachers, whether identifying central barriers or suggesting ways to improve their instructional efforts through a hypothetical "wish list", is to have their student load reduced. Aside from the barrier of content coverage pressure cited by high scorers (an issue that will be addressed elsewhere), teachers' perceptions of prominent barriers and ways to organizationally facilitate students' thinking are quite varied- -both within and between the two groups. Some data suggests however that high and lower scorers may differ in their perceptions of the effects state and AP exams have on the promotion of students' thinking.

DEPTH VS BREADTH OF CONTENT COVERAGE:

Teachers were asked a variety of questions to elicit their views on the issue of depth vs breadth of content coverage. As a whole, responses to these questions reveal more difference than similarity in the thinking of the two groups.

Both groups of teachers unanimously responded, "yes", when asked if they have a conflict between depth and breadth of content coverage. In addition, most of the high scorers (4) and a majority of lower scorers (3) view "coverage pressure" (i.e., greater breadth) as a "fairly negative" or "extremely negative" influence on the promotion of students' thinking. However, when asked an open-ended question in which they were to explain the nature of this conflict, differences emerged between the groups. Teachers' responses were categorized as indicating either an external (i.e., other than self) or internal (i.e., self) source of breadth of coverage pressure, with a few teachers indicating both (see Table 9). All of the high scorers (5) cited external sources of coverage pressure compared to only a few among lower scorers (2). For example, high scorers cited the following external sources: state exam, state or district imposed curriculum guidelines, AP course outlines and exams, students' lack of knowledge, the department head, and colleagues. The few lower scorers who did cite external

sources mentioned the AP program and district curriculum guides. Differences between high and lower scorers were minimal with respect to an internal- self source of coverage pressure, as two high scorers and one lower scorer identify self as the source of their breadth pressure.

Results consistent with the above were obtained in a later school visit when we asked teachers to rank in order of importance the following potential sources of coverage pressure: myself; department colleagues; department head; school administration; district tests or guidelines; state tests or guidelines; and, other (see Table 10). Almost all of the high scorers (4) compared to only a few lower scorers (2) identified sources other than themselves (i.e., external sources) as the primary cause for their coverage pressure. Conversely, most lower scorers (3) cited themselves as the primary cause or source for their coverage pressure compared to only one high scorer.

To review, all of the teachers say they experience a content coverage conflict. High scorers experience the conflict as primarily a situation in which external sources impose coverage demands that exceed the pace desired by the teacher, whereas lower scorers tend to impose coverage demand on themselves. This doesn't mean that high scorers do not pressure themselves toward greater coverage, only that for high scorers external sources seem to demand more coverage than the teachers themselves deem appropriate. Low scorers, however, are either in agreement with the coverage agenda of external sources or perhaps they desire even greater coverage. The findings that follow support these initial generalizations.

We also asked teachers to identify the three factors that most inhibit the promotion of students' thinking. Respondents were told to consider all facets of their working situation, including school and classroom organization, curriculum and instruction issues and materials, students, administration, etc. Almost all of the high scorers (4) identified coverage pressure (e.g., state exams, district guidelines, state guidelines) as one of the three most detrimental factors to their efforts to promote students' thinking. One high scorer went so far as to cite 2 forms of breadth pressure among his 3 most detrimental influences (i.e., a state exam, and the district's course guidelines). Only a few of the lower scorers (2) cited coverage pressure, with one of these teachers later admitting he wasn't sure what he'd do if given the time to pursue topics in greater depth! All sources of coverage pressure identified by high and lower scorers were external in origin. We also placed teachers in a hypothetical situation in which they were granted three wishes to help them promote students' higher order thinking. They were to view the question as broad ranging, and, again, the results were consistent with the above findings. Most of the high scorers (3) compared to only one lower scorer requested less content coverage as one of the three most important school

features needing modification to promote students' thinking more effectively.

Another interesting difference between the two groups can be seen in their attitudes toward the breadth oriented Advanced Placement (AP) courses. AP history courses hold students accountable for understanding a vast sweep of history for successful exam completion, though portions of the final are of an essay format which challenges students' reasoning abilities. Of the 5 teachers having had experience with AP courses, 4 were lower scorers and 1 a high scorer. All of the teachers cited immense coverage pressure in their respective AP course, yet only the high scorer felt that the AP exam "inhibited" the promotion of students' thinking. All four of the lower scorers on the other hand assessed the AP course as "supporting" the promotion of students' thinking. Similar findings emerged at the one school in the sample that participated in state testing. Like the AP format, this particular social studies state exam (and the curriculum guideline that 'drives' it) requires extensive content coverage for successful completion. The two high scorers claimed the state exam inhibits a thinking emphasis while their lower scoring colleague felt the same exam is supportive.

We also asked teachers directly to explain the nature of their coverage conflict and how they attempt to resolve it. Responses support the above findings that high scorers are more likely than lower scorers to perceive coverage pressure as externally imposed and deleterious to their teaching efforts. In addition, teachers' responses here indicate that high scorers generally attempt to resolve the conflict in the direction of less coverage compared to lower scorers by carefully selecting and prioritizing topics, ideas, and issues.

Turning first to the responses of high scorers, Harold acknowledges the coverage pressure he experiences from state guidelines and the state exam. He states:

> ...I would like to have the freedom to be flexible, you know, like today's lesson on the immigrants. It really worked. Kids were talking about what the experience would be like and what they would do in those circumstances. Some were saying that they would go home. Others felt that they would stick it out. I would like to stay with a topic like that. I could think of numerous issues I would like to pursue about this topic. But I can't....You are obligated to finish the curriculum. You also feel pressure from the chairman.

Nonetheless, Harold goes on to say:

> I do not preoccupy myself with finishing the curriculum. Instead, I attempt to teach whatever I teach well and select classroom topics and materials very carefully...I don't emphasize content coverage. It's ludicrous to attempt to cover 100 years of history in a month or two. I focus on concepts and ideas. The problem with most school courses is that they are survey courses that are homogenized.

Hanson, another high scorer, argues that teachers' attempts at broad coverage are doomed to failure since "...the amount of knowledge in the world has doubled in the last 8 years. So I see my job as enabling students to be 'good thinkers.'" He points out that the depth vs breadth conflict is really not the issue as:

> "...you could survey all of American History in a boring, non-critical, non-thinking way or you could focus for the whole 18 week semester on the causes of WWII in a boring, non-thinking non-critical way. Just going into depth doesn't necessarily mean you're teaching kids to think."

For Hanson the central issue is whether or not one emphasizes critical thinking:

> "Your commitment to promoting thinking governs your choice of curriculum...When I plan curriculum I side on the depth side of the conflict by leaving a lot of things out. I select material which is representative of particular concepts and generalizations I wish to develop. I hope that the thinking skills developed with these representative selections will be transferred to other examples... I think you can teach almost anything in a critical way, in a way that gets kids thinking - but it takes time."

Turning to the responses of lower scorers, Lisa, also experiences breadth of coverage pressure, particularly in her AP course:

> There is a comprehensive exam given the second week in May covering all the material from the 1600's to the present. This puts tremendous pressure on my classes to move through the material quickly. In our discussion of slavery we have had little time to explore the long range implications of this experience...There is simply not enough time.

Yet, when asked how she would attempt to resolve the coverage conflict if the AP course did not "impose" a curriculum of breadth, Lisa, unlike high scorers, still sides toward breadth:

> If I had to make a choice I would choose coverage because it is perhaps their only experience and probably a concluding experience with history. I really feel an obligation to at least expose them to some of the pressing issues of our time. Not to get to the 1950's and 1960's, the Cold War...is unconscionable...I think it is more important that they get exposure and that has to sacrifice depth...I want them to get as much exposure as I can, you know, shove it down their throats.

Lisa's comments reveal an internal source of coverage pressure of at least equal magnitude to the external pressure imposed by the AP curriculum guidelines. Regardless of the coverage mandates dictated by the AP course, Lisa's resolution of the breadth vs depth dilemma indicates an allegiance to exposing students to issues rather than to exploring issues with students.

Lloyd, as mentioned previously, looks to the administration for instructional guidance. Because his scope and sequence guidelines and 'skill' objectives are provided for him and accepted, he had difficulty relating to our coverage dilemma probes:

> I sort of have an idea what you're talking about but I think that issue of depth and breadth doesn't come up for me. However, he does reveal his breadth orientation (and presumably the district's) when he states; "I can't think of one area that I've taught where I'd want to go that much in depth.

Finally, we asked teachers to assess their coverage orientation with respect to particular lessons. Again differences emerge between the two groups, differences that are consistent with the above findings. Teachers were asked if their approach to the topic of the observed lesson is best characterized as oriented toward depth or breadth of coverage. All of the high scorers felt their respective lesson indicated a depth emphasis whereas all of the lower scorers viewed their lessons as breadth oriented. Yet when asked if their approach enhanced, inhibited, or had no effect on the quality of students' thinking, all of the high scorers and all but one of the lower scorers felt their approach enhanced

students' thinking![11]

A careful interpretation of the above findings suggests that lower scorers, unlike high scorers, are caught in a contradiction. That is, lower scorers make the general statement that breadth of coverage is detrimental to thinking, yet at the same time: a) claim that specific breadth-oriented lessons enhance students' thinking, and b) impose coverage pressure on themselves equal to or greater than the coverage demands articulated by the department or district. To understand lower scorers' apparent contradiction here requires a return to their instructional goals. Lower scorers tend to emphasize various forms of content acquisition as their primary instructional goal whereas high scorers view content more as a 'vehicle' to promoting thinking. For lower scorers, thinking is incorporated into their primary mission of teaching content, whereas for high scorers thinking is the central focus, with content understanding as a necessary component. Since lower scorers place greater emphasis on content acquisition it is not surprising that compared to high scorers they tend more toward 'breadth' of coverage. This would explain lower scorers' tendency to self-impose coverage demands equal to or greater than the coverage demands of their social studies department or school district. (This may also explain their positive assessment of breadth-oriented Advanced Placement courses).

It is a bit more difficult to explain lower scorers' belief that in general breadth of coverage is detrimental to a thinking emphasis when at the same time they feel that their own specific lessons oriented toward breadth promote thinking. It is possible, however, for lower scorers to be consistent here. They could believe that their breadth-oriented lessons promote thinking but not as much as a depth-oriented lesson. Their 'breadth lesson' would then be consistent with their overriding primary goal to expose students to many facts and ideas in the social studies. Unfortunately, a follow-up probe was not attempted with respect to this issue. However, the dissatisfaction some of the lower scorers expressed about their breadth-oriented lesson during a post-observation interview, and the almost apologetic tenor of its expression, suggests that lower scorers have not resolved their coverage dilemma in the finely reasoned manner constructed above. Rather lower scorers seem to vacillate between teaching for content

---

[11]Due to illness or snow days 4 of 5 high scorers and 3 of 5 lower scorers were asked coverage questions relating to a specific lesson observation. This is unfortunate as the 2 lower scorers not asked were most candid about their breadth orientation, making it very likely that 5 of 5 lower scorers would have described their lesson as characterized by breadth. Whether these two teachers would see their respective lesson as enhancing or inhibiting the quality of students' thinking is less clear.

coverage that is both self and externally imposed, and a desire to engage students in thinking and more substantial analysis of topics which requires greater depth of coverage. Further research is clearly necessary on this matter.

REFLECTIONS ON STUDENTS:

High and lower scoring teachers unanimously agree that at times students resist school work that demands thinking. The types of thinking tasks teachers identify as most frequently resisted by the less receptive students are extremely varied and difficult to classify in ways that might distinguish between high and lower scorers. Of the 20 or more thinking tasks identified by teachers, in only two instances did two teachers from either group identify the same task as most frequently resisted by less receptive students. The conclusion to be drawn here is that teachers perceive students as resisting a wide variety of tasks requiring thinking.[12]

Both groups of teachers unanimously agree that some students are more receptive than others to assignments, activities, and tasks requiring thinking. Using a set of structured questionnaire items teachers were asked to reflect not only on factors leading some students to perform "above average" on thinking tasks, but also to reflect on factors that influence "below average" performance. We were curious to see if teachers emphasized different factors with above and below average 'student-thinkers'. Teachers were to rate on a 5-point scale (1= no influence to 5= major influence) the degree to which they believe each of 12 factors play a role in the performance of above and below average 'student-thinkers.[13] A mean score was computed for both groups of teachers on each of the 12 factors. The 3 factors receiving greatest emphasis (i.e., the highest mean scores) by each group (see Tables 11 and 12) were compared and will be discussed below.

The data suggests that high and lower scoring teachers agree upon the most important factors for student performance on thinking tasks. Both groups emphasize the same 3 factors (out of a possible 12) when attempting to explain the performance of above average student-thinkers, and 2 of 3 factors with respect to below average

---

[12]However, teachers' responses here do suggest that the tasks high scorers pose to less receptive students are more readily identifiable as tasks requiring thinking than are the tasks posed by lower scorers. This will be discussed later in this section.

[13]The 12 factors are peer group pressure, family background/situation, students' employment responsibilities, students' experiences in prior social studies courses, students' rapport with teacher, students' inherent cognitive capacity, students' learned thinking skills/abilities, students' present knowledge, students' writing abilities, students' level of self-confidence and self-esteem, students' motivation, and any other factor not mentioned above.

student-thinkers. The three factors both groups emphasize with above average student-thinkers are "learned thinking skills/abilities", "cognitive capacity", and "motivation".[14] The two factors both groups of teachers emphasize with below average student-thinkers are (a lack of) "motivation" and (lack of) "self-confidence/esteem".

Second, high and lower scoring teachers do not emphasize the importance of "cognitive capacity" with below average student-thinkers but do so with respect to above average students. Both groups apparently feel that factors other than cognitive capacity best explain the inability of low performing students to exhibit at least average if not above average performance on thinking tasks.

Third, "motivation" is the only factor to be emphasized by both groups of teachers with respect to both types of students. Teachers may intuitively know what educational researchers are now only beginning to suggest (Dweck, 1986); that is, the development of good thinkers most fundamentally involves changing students' general attitudes and dispositions rather than only expanding students' knowledge structures, improving cognitive processes, or developing specific thinking skills.

Fourth, the relatively high mean score across all 12 factors (i.e., 3.81 or 3.83) suggests that both groups of teachers believe a variety of factors play a role in students' performance on thinking tasks. Teachers and not just researchers realize the complexity of the instructional/learning/thinking process.

Possibly more important than teacher perceptions of factors that influence student performance on thinking tasks is the teacher's sense of influence on these factors. We suspect that the more a teacher believes he/she can influence factors associated with thinking, the higher will be the teacher's expectations for student performance.

Teachers were asked to rate on a 5-point scale (1= no influence to 5= major influence) the degree to which they believe they can influence in one year 7 factors contributing to students' performance on thinking tasks.[15] A mean score was computed for

---

[14]Due to a three-way tie in mean score on the third most important factor cited by high scorers, 5 rather than 3 factors are listed. The two other factors emphasized by high scorers are "self- confidence/esteem" and "teacher-student rapport."

[15]The 7 factors are: students' rapport with teacher, students' learned thinking skills/abilities, students' present knowledge, students' writing abilities, students' level of self-confidence and self-esteem, students' motivation, and 'other'.

both the high and lower scoring groups on each of the 7 factors. The 3 factors highlighted by each group as being most amenable to positive influence (see Tables 13 and 14) were compared and will be discussed below. The data suggests that high and lower scoring teachers hold very similair views regarding the factors they believe they can most influence with students in the course of a year. Out of 7 possible performance factors both groups believe they can most influence the same 3 factors with respect to above average student-thinkers, and 2 of 3 factors with respect to below average student-thinkers. The three factors both groups emphasize with above average student-thinkers are "students' knowledge base", "teacher-student rapport", and "students' writing skills".[16] The two factors both groups of teachers emphasize with below average student-thinkers are "students' thinking skills/abilities" and "teacher-student rapport". Only one factor was consistently ranked as highly amenable to teacher influence among high and lower scorers, regardless of student type. This was "teacher-student rapport."

We found that lower scorers believe they can influence the performance of above average student-thinkers more than high scorers, for the mean score of lower scorers across 7 factors is 3.69 vs 3.20 for high scorers. But this conclusion is suspect, because a number of high scorers during interviews mentioned that they felt little need to influence or improve some factors with above average student-thinkers; these students were already outstanding in such areas as motivation and teacher-student rapport.

In contrast, high scorers believe they can influence the performance of below average student-thinkers more than lower scorers. Four pieces of evidence support this conclusion. First, the mean score for high scorers across all 7 performance factors is 3.61 vs 3.08 for lower scorers, a difference of over .5 per factor. Second, in a direct probe of this issue, teachers were asked whether or not it is more difficult to emphasize thinking with low achieving students.[17] Most high scorers (4), but only a few lower scorers (2), said it was <u>not</u> more difficult to emphasize thinking with low achieving students. For example, Hilary, a high scorer, says she "stays with a student" who responds "I don't know" to an opinion, analysis, or evaluation question. "I let them know at

---

[16]Due to a two-way tie in mean score for the third factor among high scorers, 4 rather than 3 factors are listed. The one other factor emphasized by high scorers is "thinking skills/abilities."

[17]An assumption is made here that when teachers think about their low achieving students and their below average thinkers they have, essentially, the same population of students in mind.

the beginning of the year that an 'I don't know' answer is unacceptable, that I will not go away until they get it right." Hilary feels that "prior experience, lack of confidence, and mental laziness" have combined to make "I don't know" a safe response, one that "will make the teacher seek a response elsewhere." However, "after a few successes they want to be called on and try again."

A third strand of evidence comes from an interview in which teachers were asked to explain the lack of involvement on the part of some students in a lesson recently observed by the researcher.[18] All of the lower scoring teachers (5) blamed students themselves for lack of lesson involvement while only a few high scoring teachers (2) did. Lower scorers suggested that students don't like to talk, don't feel like working, aren't interested, haven't done the reading, are lazy, fear ridicule, cannot articulate their thoughts, etc. The few high scorers who blamed students mentioned the following three student deficiencies: some have personal problems, are afraid to risk their ideas publicly, and, don't have a clue as to how to make a prediction/hypothesis. Recall that of the classrooms observed, most lower scorers (3) had classes composed primarily of high achieving students, whereas most high scorers (4) had either a diverse range of achievers or predominantly low achievers. Yet one finds the lower scorers blaming students for lack of lesson involvement, not the high scorers.

A fourth piece of evidence suggesting that high scorers feel more influential with below average student-thinkers is that a majority of high scorers (3) but no lower scorers placed partial blame on themselves for disengagement among some students during a specific lesson. High scorers incriminated themselves in the following ways: hadn't learned the names of all the students and therefore had to impersonally point his finger to get responses, or avoided calling on some students; had kids continue to give their opinions even though responses were redundant which led some students to tune-out; calling on kids row by row rather than randomly which led some to tune out until it was their turn to be called on; and, needed to spend more time with particular students on their responses. Self-attributions by high scorer here suggests greater locus of control on the part of these teachers; the responsibility for promoting thinking seems to rest squarely with the teacher, not with the student.

To summarize this section on teachers' perceptions of students, much agreement exists between high and lower scoring teachers. Both groups unanimously agree that students at times

---

[18]An assumption is being made here that the students who were not engaged in the lesson are also, in large part, the students who perform below average on thinking tasks.

resist thinking tasks, and that some students are more receptive than others to tasks requiring thinking. Both groups emphasize the same factors (3 of 3) as most important for the performance of above average student-thinkers, and the same factors (2 of 3) with respect to below average student-thinkers (see Table 15). Both groups are also in agreement regarding the thinking performance factors they feel they can most influence with students. High and lower scoring teachers believe they can most influence the same factors (3 of 3) with respect to above average student-thinkers, and the same factors (2 of 3) with respect to below average student-thinkers (see Table 16).

The two groups seem to part ways, however, in the degree to which they believe they can influence the performance levels of below-average student-thinkers. High scorers compared to lower scorers are more likely to say that it is not more difficult to emphasize thinking with low achieving students, and blame themselves rather than students for lack of student lesson involvement.

CONCEPTIONS OF THINKING:

Since prior research shows that many of the ideas guiding teachers' practices are implicitly understood and are therefore difficult for teachers to articulate (Clark & Peterson, 1987), we used several questions to elicit information on teachers' conceptions of thinking. The two most direct questions were:

> -Do you have a conception of thinking that guides your teaching? If so, summarize its main aspects.
>
> -Consider your best thinkers, what distinguishes them from your other students?

The first question asks teachers directly to explain their conception. The second question, an indirect probe, assumes that a teacher's conception of the good thinker is closely related to his/her conception of thinking. Teacher responses to several other questionnaire items that further illuminate teachers' conceptions of thinking were also included in the analysis.

The framework developed to help analyze teachers' conceptions of thinking draws upon major distinctions made in some of the scholarly literature on thinking, specifically thinking as dispositions, as content mastery, or as skills. With regard to dispositions, Schrag (1987) has argued that thinking is closely linked to the character traits of reflectiveness and mental flexibility. Other dispositions frequently associated with thoughtfulness include: curiosity to question and explore; an insistence that claims be supported by reasons (and that the reasons themselves be scrutinized); confidence in one's thinking;

motivation; and, a willingness to view an issue or problem from different perspectives. Here are a few example-excerpts of dispositional language taken from teachers' conceptions of thinking:

> -A good thinker isn't afraid if someone challenges a position. A good thinker is willing to take a look at someone else's hypothesis or theory even if it's 180 degrees apart from his own, rather than a dogmatic knower.
>
> -I want to develop in students an attitude that nothing is finally closed...If that's skeptical I don't think it's skeptical in a negative sense.

The second component or perspective on thinking highlights the importance of knowledge of subject matter understanding for effective thinking, as suggested, for example, in recent research on expert/novice problem-solving and various information processing models of thinking and learning. According to this perspective, one cannot think in a content vacuum, that is, one must have something to think about. Effective thinking about a problem or task occurs only when one possesses sophisticated understanding of the content to which one's thinking is directed, be it elevator repair or open-heart surgery. Teachers also emphasized content understanding; for example:

> -You build off a knowledge base which is essential.
>
> -You can only think when they have enough info to get a detailed picture of an event.

Finally, the skills perspective defines thinking in terms that usually transcend specific subject matter. A good thinker, for example, can detect bias, identify a problem, muster evidence, and analyze or evaluate a body of material, etc. Skills can be construed as domain- specific, for example, legal reasoning in law, or solving proofs in geometry. In short, the skills view states that students must possess a variety of skills, techniques, strategies, and heuristics when approaching a problem or task. Teachers also used a skills language of thinking;

> -Thinking is determining validity based on reasoning, intuition, and emotional skills.
>
> -There are five process skills; recall, organization, analysis, synthesis and evaluation.

Simply counting teachers' responses to various questionnaire items revealed little difference in the frequency with which high and lower scorers use dispositional, skill, or content language to describe their conceptions of thinking. For example, when teachers were asked to describe their conception of thinking, high scorers primarily used skill (5) and dispositional (3) language (see Table 17). Low scorers' responses are very similar; that is, skills (4) and dispositional (5) language predominate, with some teachers also referring to content (3). Similarly, when teachers were asked to explain how observed lessons (2-3 lessons per teacher) promoted students' thinking (see Table 18), the responses of high and lower scorers again revealed little difference. Based on a total of 13 lessons from each group of teachers, all high and lower scorers (5) used content language, all high (5) and most lower scorers (4) used skills language, while only one teacher from either group used dispositional language.[19]

Though both groups respond in like fashion to the two different probes into teachers' conceptions of thinking, notice that there is a shift in language usage. The greater emphasis placed on content language when discussing lessons is probably due to the fact that teachers are not talking here in general terms, but rather about specific thinking tasks with respect to specific content. The lack of dispositional language is not surprising, one probably wouldn't expect a teacher to emphasize the development of thoughtful dispositions in a time frame of one lesson. Whatever the actual reason(s) for the shift in emphasis, one can conclude that high and lower scorers use dispositional, skill, and content language in roughly equal measure when talking about thinking generally or with respect to specific lessons.

Significant differences begin to emerge between the two groups, however, when teachers' conceptions are viewed more qualitatively and in their entirety; that is, high scorers' conceptions of thinking tend to be lengthier, more detailed, and more elaborate than the conceptions of lower scorers. The thinking conception of one teacher from each group will be described to highlight these differences. The high and lower scoring teacher to be described here possess neither the most nor least developed conception of thinking in their respective group.

Hilary, a high scorer, describes her conception primarily in skill and dispositional language. She mentioned skills that include the ability to "analyze", "focus on an argument and support it", "determine cause and effect", "induction", "evaluating, based upon supportable evidence", "determining point

---

[19]Skill language was used by high scorers on 10 of 13 lessons and by lower scorers on 7 of 13 lessons. Content language was used by high scorers on all 13 lessons and by lower scorers on 11 of 13 lessons.

of view and identifying its effects", "see relationships", and "draw conclusions." She emphasizes that "the teacher must give students the opportunity to develop these abilities."

In addition, Hilary subscribes to the Grandville High School skills model of thinking because she's "in the department and it's so central to the program." This model combines the work of Bloom, and Beyer's skills approach to thinking. Hilary is well-versed in the specifics of the model, as she quickly identified the component thinking skills (i.e., recall, organization, analysis, synthesis, evaluation, and divergent thinking) and then discussed their hierarchical nature; "The 'higher' skills entail the skills below...organization involves recall, analysis also involves organization and recall, synthesis involves analysis...Divergent thinking is at the top and is the hardest for students to do."

Though Hilary is versed in skills language, she also emphasizes important elements of thinking with dispositional language. In fact, one of the most important attributes of the good thinker for Hilary is "intellectual curiosity." She emphasizes that "students who perform best on thinking tasks often have an intellectual curiosity...not necessarily equated with highest 'inherent cognitive capacity'." Hilary encourages students to ask questions, and ultimately wants them "to see that question asking is a sign of intelligence not ignorance; to get them to know what they don't know and want to find out." She finds the attribute of intellectual curiosity to exist not only at the individual level, but also at the classroom level. It is this 'type' of student, and class, that she most enjoys.

As discussed earlier with respect to Hilary's goals, she places great emphasis on developing students' appreciation of and willingness to experience alternative perspectives. Above the blackboard on the front wall is a poster stating the following: "Where you stand depends on where you sit." This attitude, or general understanding or belief in the importance of perspective is, according to Hilary, "essential for students' thinking." With great elaboration she describes lessons she uses to bring this point or 'viewpoint' across to students. In fact, 'perspective taking' is the topic of Hilary's very first lesson each semester and provides the central theme for her courses.

Consistent with her emphasis on perspective taking and intellectual curiosity, Hilary tries always to keep the following in mind when leading a discussion: "Never ask a question where you already know what you want for an answer." Another disposition she finds important for thinking is a "willingness to take risks".

When discussing how specific lessons challenge students to think Hilary states that she much prefers to talk about her efforts with respect to the actual content of the lesson.

Curiously, Hilary claims she lacks "the cognitive scientist's language to describe thinking" (except for the department's 6-skill thinking model), and doesn't consciously apply it when lesson planning or engaged in classroom instruction. Instead, she "intuitively knows what questions are 'hard', 'harder', and 'hardest'." However, her quite elaborate conception of thinking, combined with references to specific classroom behaviors she employs to facilitate students' thinking, may belie the above comments minimizing the importance of a thinking conception for her instructional success.

Lloyd, like most of his fellow lower scorers, provides a fairly brief conception of thinking, one that implies concern for dispositions and skills. Lloyd highlights several skills: "ability to connect seemingly unrelated ideas", "ability to communicate", "categorize or group things", "use application skills", and "interpretation, categorizing, sequencing...analysis...and evaluate the facts". Lloyd mentions that he has read Norris Sanders', Classroom Questions, and finds that "it is useful because you can thumb through the book for his hierarchy and see some applications of the questioning techniques." Though Lloyd did not summarize Sander's hierarchy, it is reasonable to assume he knows, because many of the skills he mentions are also used by Sanders (i.e., interpretation, application analysis, evaluation). Lloyd's concern for dispositions emerges when he defines intelligence as "curiosity and eagerness and ability to connect different ideas."

To briefly summarize this section on teachers' conceptions of thinking, high and lower scorers tend to use content, skill, and dispositional language in equal measure when discussing their conception of thinking generally or with respect to specific lesson efforts. Important differences begin to emerge however when teachers' conceptions of thinking are viewed more qualitatively. That is, most high scorers compared to only one lower scorer offered a conception of thinking that was fairly detailed, lengthy, and elaborate.[20]

CONCLUSION:

This exploratory study of 5 areas of teachers' thought (i.e., goals, barriers, depth vs breadth, perceptions of students, and conceptions of thinking) found both differences and similarities in the thinking of teachers scoring high or low in their promotion of

---

[20]Space limitations prevent the presentation of additional examples to help substantiate this point. However, with respect to length, summaries of high scorers' conceptions of thinking averaged 47 lines of text whereas lower scorers' summaries averaged 19 lines. See Onosko (1988) for complete summaries of each teacher's conception of thinking.

thoughtful classroom discourse. Individual questionnaire items revealed neither complete agreement within the high or lower scoring groups nor dramatic differences between them. But when results of related questions were combined, some consistent patterns emerged that indicate areas of similarity and difference between groups. Both high and lower scorers hold similar views on barriers to promoting students' thinking, specifically large class size and total student load. Unfortunately, neither group suggested strategies to minimize the effect of these barriers. Instead, we found a pervading pessimism about change, especially with respect to student load.

Many similarities in the thinking of the two groups also emerged with respect to students, especially in the factors perceived to be most important for successful performance on thinking tasks, and the factors that teachers believe they can most influence with students. However, various strands of evidence suggest that high scorers compared to their lower scoring counterparts feel more efficacious when it comes to promoting the thinking of below-average student-thinkers.

Important differences seem to exist between the two groups with respect to their instructional goals and their views on the issue of depth vs breadth of content coverage. Lower scorers tend to emphasize content acquisition as their primary instructional goal whereas high scorers view content more as a 'vehicle' to promoting thinking. Stated another way, thinking is incorporated into lower scorers' content mission, whereas high scorers place thinking as the central focus with content understanding a valued outcome. Lower scorers prefer to <u>expose</u> students to ideas and issues, whereas high scorers prefer to <u>explore</u> ideas and issues with students in depth. This emphasis on thinking and content exploration leads high scorers to reduce content coverage compared to lower scorers, and to voice greater objection to external sources of coverage pressure.

With respect to teachers' conceptions of thinking, high scorers offered lengthier, more detailed, and more elaborate conceptions than lower scorers, though some of the professional language used by each group was very similar (i.e., content, skills, or dispositions). It is interesting to note that the group possessing more elaborate conceptions of thinking also place greater emphasis on thinking as an instructional goal and on depth of coverage to better insure thoughtful exploration of subject matter. But the exact relationship between teachers' conceptions of thinking, and their instructional goals and resolution of the coverage dilemma is unclear.

Acknowledging, again, the exploratory nature of this study, four recommendations for teacher education can be suggested from conclusions drawn above. Teacher educators need to find ways to:

a) increase teachers' commitment to and rationale for promoting students' thinking as a primary instructional goal;

b) enhance teachers' conceptual understanding of thinking;

c) increase teachers' sense of influence with low achieving students by changing teachers' perceptions of these students; and

d) increase teachers' understanding of the positive relationship between depth of coverage and the promotion of thinking

REFERENCES

Adler, M. (1982). The paideia proposal: An educational manifesto. New York: Macmillan.

Boyer, E. (1983). High School: A report on secondary education in America. New York: Harper & Row.

Brophy, J. & Rohrkemper, M. (1981). The influence of problem ownership on teachers' perceptions of and strategies for coping with problem students. Journal of Education Psychology, 73, 295-311.

Clark, C. & Peterson, P. (1987). Teacher thought processes. In M. Wittrock (ed.), Handbook of Research on Teaching: 3rd Edition. New York: Macmillan.

College Board. (1983). Academic preparation for college: What students need to know and be able to do. New York: Author.

Cuban, L. (1984). How teachers taught: Constancy and change in American Classrooms. New York.

Cuban, L. (1984b). Policy and research dilemmas in the teaching of reasoning: Unplanned designs. Review of Educational Research, 54, 655-681.

Dweck, C. (1986). Motivational processes affecting learning. American Psychologist, 41 (10), 1040-1048.

Good, T. (1987). Two decades of research on teacher expectations: Findings and future directions. Journal of Teacher Education, July-August, 32-47.

Goodlad, J. (1984). A place called school: Prospects for the future. New York: McGraw-Hill.

Hoether, J. & Ahlbrad, W. (1969). The persistence of recitation. American Educational Research Journal, 6 (2), 145-167.

King, L. (1980). Student thought processes and the expectancy effect (Research Report No. 80-1-8). Churchlands, Perth, Australia: Churchlands College of Advanced Education.

Lightfoot, S. (1983). The good high school: Portraits of character and culture. New York: Basic Books.

Marland, P. (1977). A study of teachers' interactive thoughts. Unpublished doctoral dissertation. University of Alberta, Edmonton, Canada.

Munby, M. (1983). A qualitative study of teachers' beliefs and principles. Paper presented at the annual meeting of the American Educational Researchers Association, Montreal.

McNeil, L. (1983). Defensive teaching and classroom control. In M. Apple & L. Weis (Eds.), Ideology and practice in schools (pp. 114-143). Philadelphia: Temple University Press.

Newmann, F. (1988a). Higher order thinking in the teaching of social studies: Connections between theory and practice. In Informal reasoning and education, D. Perkins, J. Segal, & J. Voss (Eds.). Hillsdale, NJ: Erlbaum.

Newmann, F. (1988b). Can depth replace coverage in the high school curriculum? Phi Delta Kappan, Feb., pp. 345-348.

Page, R. (1984). Perspectives and processes: The negotiation of educational meanings in high school classes for academically unsuccessful students. Unpublished doctoral dissertation, University of Wisconsin, Madison.

Peterson, P. & Barger, S. (1984). Attribution theory and teacher expectancy. In J.B. Dusek (Ed.), Teacher expectancies (pp. 159-184). Hillsdale, NJ: Erlbaum.

Perkins, D.N. (1986). Knowledge as Design. Hillsdale, NJ: Lawrence Erlbaum & Associates.

Perrone, V. & Associates (1985). Portraits of high schools: A supplement to high school: A report on secondary education in America. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Powell, A., Farrar, E., & Cohen, D. (1986). The shopping mall high school. Boston, MA: Houghton Mifflin.

Rutter, R. (1986). Profile of the Profession. Social Education, 50 (4), 252-255.

Schrag, F. (1988). Thinking in School and Society. New York: Rutledge & Kegan Paul.

Schulman, L. (1987). Knowledge and teaching: Foundations of a new reform. Harvard Educational Review.

Shavelson, R. & Stern, P. (1982). Research on teachers' pedagogical thoughts, judgments, decision, and behavior. Review of Educational Research, 51, 455-498.

Siegel, H. (1988). Educating Reason. New York: Rutledge & Kegan Paul.

Sirotnik, K. (1983). What you see is what you get: Constancy, persistency, and mediocrity in classrooms. Harvard Educational Review, 53 (1), 16-31.

Sizer, T. (1984). Horace's compromise: The dilemma of the American high school.. Boston: Houghton Mifflin.

Swartz, R. (1987). Thoughtfulness: Is H.S. a place for thinking? In the Center Newsletter, 2(1). Madison, WI: National Center on Effective Secondary Schools.

Zeichner, K., Teitelbaum, K. (1982). Personalized and inquiry-oriented teacher education. Journal of Education for Teaching,8, 95-117.

Table 1

Comparison of Background Characteristics and Organizational Dimensions

	High Scorers	Lower Scorers
--Sex: Male	4	4
Female	1	1
--Years teaching social studies	14.6	17.0
--Education: Masters' degree	1	2
Masters' plus credits	4	2
Doctorate	0	1
--Job satisfaction: Most of time	4	4
Some of time	1	1
--Subject of course observed:		
History	3	4
Social Science	2	0
Social Studies	0	1
--Number of schools represented	3	4
--Teacher load: Teaching periods per day	4.6	3.6
Average class size	28.6	26.4
Total number of students	131.8	97.6
Number of course preparations/day	2.2	2.4
--Class composition: % Black	12.6	28.8
% Asian	5.0	1.2
% Hispanic	11.0	12.8
--Prior social studies achievement of students in the observed class:		
High achievers	1	3
Diverse range of achievers	2	1
Low achievers	2	1

TABLE 2
HIGHEST PRIORITY GOAL

	High Scorers	Lower Scorers
-develop critical thinking and problem solving abilities	4	2
-teach facts, concepts, and theories of history and the social sciences	1	2
-teach past and present problems and issues faced by the U.S. and World	0	1
-teach social values and foster citizenship	1	1
-develop creative thinking abilities	0	0
-develop reading and writing abilities	1	0
-develop discussion skills	0	0
-teach students how to study, take notes and learn	0	0
-develop self-confidence and self-esteem	0	0
-other	0	0
TOTAL	7*	6*

*Note: some teachers cited more than one highest goal.

TABLE 3
GOALS THAT FOCUS LESSON PLANNING

	High Scorers	Lower Scorers
-develop critical thinking and problem solving abilities	5	3
-teach facts, concepts, and theories of history and the social sciences	3	4
-teach past and present problems and issues faced by the U.S. and World	0	0
-teach social values and foster citizenship	4	3
-develop reading and writing abilities	2	2
-develop creative thinking abilities	0	0
-develop discussion skills	0	0
-teach students how to study, take notes and learn	1	0
-develop self-confidence and self-esteem	0	0
-other	0	0

TABLE 4
TEACHER INTEREST

	High Scorers	Lower Scorers
Would you say that in general exposing to subject matter content is:		
a) more	0	1
b) equally	0	3
c) less	5	1
interesting for you as a teacher, as developing students' thought and reasoning processes?		

Table 5

Teacher Satisfaction

DERIVED FROM:	High Scorers	Lower Scorers
--Students Thinking	5	1
--Students Responding	1	5
--Working with Colleagues	3	2
--Lesson Planning	1	1

Table 6

Teacher Assessment of Potential Barriers

	Extremely Negative 1	2	3	4	Extremely Positive 5
	High Scorers		X	Lower Scorers	
--Large total # of students	1.8			1.2	
--Large class size	1.8			1.6	
--Large student course load	2.4			2.4	
--Short 45 min. class period	2.8			3.0	
--Large # of graduation requirements	3.0			2.8	
--4-5 class sessions per week	3.6			3.6	
--Student ability grouping	2.8			3.8	

Table 7

Three Most Inhibiting Barriers
(# of teachers citing a particular barrier)

	High Scorers	Lower Scorers
--Large number of students (load)	3	4
--Large class size	3	3
--Content coverage pressure	5*	2
--Lack of planning time	1	2
--Short 45 min. class periods	1	1
--Student ability grouping	1	0
--Large # of classes per day	1	0
--Diversity of students' abilities	0	1
--4-5 class sessions per week	0	1
--Lack of quality lesson materials	0	1
	15	15

*The actual number of high scorers citing coverage pressure is 4, as one high scorer mentioned coverage pressure twice (i.e., district course-guideline coverage pressure and coverage pressure resulting from the state exam).

Table 8

Three Wishes to Promote Thinking

	High Scorers	Lower Scorers
--Fewer students	4	4
--Freedom to cover less content	3	1
--Better curriculum materials	1	3
--Staff development to improve my teaching	2	2
--Paid leave of absence to think, read, or develop curriculum	2	1
--Team teaching	0	1
--Team/group planning	0	1
--More lesson planning time	0	1
--Peer observation	1	0
--Fewer classes per day	1	0
--Different group of students	1	1

Table 9

Nature of Coverage Conflict

	High Scorers	*Lower Scorers
External source (breadth pressure)	5	2
Internal source (breadth pressure)	2	1

*The responses of two lower scorers could not be categorization as either external or internal. Further probing revealed this was due to the fact that both teachers are from a school participating in a 'top-down' district-wide program of curriculum revision. The two teachers receive and accept the district's curriculum mandates and therefore experience much less the dilemma of content coverage inherent in planning efforts. See the response of "Lloyd" later in this section for details.

Table 10

Primary Source of Breadth of Coverage Pressure

	High Scorers	Lower Scorers
Internal		
--myself	1	3
External		
--Department colleagues	0	0
--Department head	1	0
--School administration	0	0
--District tests or guidelines	1	1
--State tests or guidelines	3	0
--Other: AP Exam	0	1
Total	6*	5

*One high scorer cited 2 primary sources of pressure (i.e., district guidelines and state tests).

Table 11

Teachers' Perceptions of the Three Most Important Factors for the Performance of Above Average Students on Thinking Tasks

High Scoring Teachers	Lower Scoring Teachers
Confidence/self-esteem (4.6)	Cognitive capacity (4.6)
Motivation (4.6)	Motivation (4.4)
Learned thinking abilities (4.2)	Learned thinking abilities (4.4)
Cognitive capacity (4.2)	
Teacher-student rapport (4.2)	

Note: The mean score across all 12 factors was quite similar for both groups of teachers; 3.81 for high scorers and 3.83 for lower scorers. The difference in mean score between the two groups of teachers on any given factor did not exceed .6, except "self-confidence/esteem" where the mean for high scorers was 4.6 vs 3.6 for lower scorers.

Table 12

Teachers' Perceptions of the Three Most Important Factors for the Performance of Below Average Students on Thinking Tasks

High Scoring Teachers	Lower Scoring Teachers
Learned thinking abilities (4.6)	Motivation (4.6)
Confidence/self esteem (4.6)	Confidence/self-Esteem (4.2)
Motivation (4.4)	Teacher-student rapport (4.2)

Note: The mean score across all 12 factors was identical for both groups of teachers; 3.83 for high scorers and 3.83 for lower scorers. The difference in mean score between the two groups of teachers on any given factor did not exceed .6.

Table 13

Teachers' Perceptions of the Three Factors of Performance They Can Most Influence With Above Average Student-Thinkers

High Scoring Teachers	Lower Scoring Teachers
Student knowledge base (3.8)	Student knowledge base (4.2)
Thinking skills/abilities (3.6)	Teacher-student rapport (4.2)
Teacher-student rapport (3.4)	Writing skills (4.2)
Writing skills (3.4)	

Note: The mean score across all 7 factors was 3.20 for high scorers and 3.69 for lower scorers, or an average of about .5 higher for lower scorers on each of the 7 factors.

Table 14

Teachers' Perceptions of the Three Factors of Performance They Can Most Influence With Below Average Student-Thinkers

High Scoring Teachers	Lower Scoring Teachers
Teacher-student rapport (4.2)	Teacher-student rapport (3.8)
Student knowledge base (4.0)	Self-confidence/esteem (3.4)
Thinking skills/abilities (3.8)	Thinking skills/abilities (3.4)

Note: The mean score across all 12 factors was 3.61 for high scorers and 3.08 for lower scorers, or an average of a little more than .5 higher for high scorers on each of the 7 factors.

Table 15

Factors High and Lower Scorers Agree Are Most Important For Student Performance on Thinking Tasks

Above Average Student-Thinkers
--thinking skills/abilities
--inherent cognitive capacity
--motivation

Below Average Student-Thinkers
--self-confidence/self-esteem
--motivation

Table 16

Factors High and Lower Scorers Agree They Can Most Influence With Students

Above Average Student-Thinkers
--knowledge base
--teacher/student rapport
--writing skills

Below Average Student-Thinkers
--thinking skills/abilities
--teacher/student rapport

Table 17

Breakdown of the Number of Teachers Using Dispositional, Skill, and Content Language to Describe Their Conception of Thinking

	High Scorers	Lower Scorers
--Dispositional	3	5
--Skill	5	4
--Content understanding	1	3

Table 18

Breakdown of the Number of Teachers Using Dispositional, Skill, and Content Language to Discuss How Their Lessons Promote Thinking

	High Scorers	Lower Scorers
--Dispositional	0	1
--Skill	5	4
--Content understanding	5	5

IV

CLASSROOM PRACTICES OF HIGH AND LOWER SCORING TEACHERS

Joe Onosko

This chapter presents findings on the classroom practices of high and lower scoring teachers at promoting students' thinking through thoughtful classroom discourse. The intent here is to better understand the ways in which the classroom practices of the two groups are similar and different. Comparisons are reported between the two groups on 10 dimensions of thoughtfulness using the observational scales described earlier. Frequencies are also reported on certain generic teaching practices such as the type(s) of: a) reading materials used; b) writing tasks required; and, c) instructional formats employed during class. This information will be supplemented and enriched in two ways. First, examples of observed lessons with a high and a low score on each of the 10 dimensions of practice are described. And four entire lessons (2 from each group) are summarized to provide a better understanding of differences that exist between the two groups of teachers across multiple dimensions of thoughtful practice.

COMPARING HIGH AND LOWER SCORERS ON 10 DIMENSIONS OF THOUGHTFULNESS:

High and lower scorers can be compared across three different lesson sets: a) all lessons; b) teachers' most thoughtful or best lessons; and, c) teachers' less thoughtful or remaining lessons. "All lessons" refers to the 9 classroom observations of each teacher over the course of the 3 site-visits, or 45 lessons per group.[1] A subset of all lessons are teachers' "best" or most thoughtful lessons.[2] These lessons were used to select teachers as either high or lower scorers at promoting thoughtful classroom discourse. Teachers' "remaining" or less thoughtful lessons refers to those lessons not used to select teachers as either high or lower scorers. Because the findings reveal that with respect to teachers' remaining lessons sizeable differences exist on the exact same scales as that found

---

[1] Due to illness, snow day, etc., one high and one lower scorer from each group was observed on only 8 occasions, reducing the lesson sample from 45 to 44 for each group.

[2] The teachers with the greatest number of lessons receiving a high score of "4" or "5" on the six scales representing the minimal requirements of a thoughtful lesson (i.e., HOTAV) were designated high scorers (except for "giving students time to think" which required a score of "5"). High scorers' lessons here constitute their best lessons. Lower scorers, due to the fact that none of their lessons met the criteria used to select high scorers, comprised the 5 teachers with the fewest number of lessons receiving a moderately high score of "3" or more on each of the minimal requirement scales (except for "time to think" which required a score of "4"). Lower scorers lessons here constitute their best lessons.

across all lessons [3], the comparison of high and lower scorers will focus on a) all lessons and b) best lessons.

With respect to teachers' best lessons and the first 6 scales constituting the minimal requirements of a thoughtful lesson, our primary interest is not whether differences exist between the two groups, but whether the differences are sizeable or meaningfully different. High scorers, by definition, will outscore lower scorers on these 6 scales because the initial identification of teachers as high or lower scorers was based on their respective performance on the 6 minimal requirement scales.[4] Because teachers' best lessons are included in the set of all lessons, high scorers are also likely to outscore lower scorers across all lessons on the above 6 dimensions. Therefore, our primary interest is whether or not the magnitudes in favor of high scorers are sizeable or educationally significant. Teachers will also be compared on the remaining 4 dimensions of thoughtful practice. Because these dimensions were not used to initially identify teachers as high or lower scorers, it remains to be seen what group differences might emerge on these scales, and whether these differences are sizeable or educationally significant. Finally, we will check to see if high scorers consistently facilitate thoughtful discourse more so than lower scorers. That is, are differences observed between the two groups on teachers' best lessons maintained across all lessons?

To summarize, our analysis of the 10 dimensions of thoughtfulness can be stated in the form of three questions:

1) With respect to all lessons and teachers' best lessons are there sizeable differences between high and lower scorers on the first 6 dimensions of practice constituting the minimal requirements of a thoughtful lesson?

2) With respect to all lessons and teachers' best lessons how do high and lower scorers compare on the remaining four dimensions of thoughtful practice? Are there sizeable differences, and if so, in favor of which group?

3) Do high scorers consistently facilitate thoughtful discourse more so than lower scorers; that is, are there sizeable differences between the two groups regardless of lesson set analyzed?

---

[3] Comparing all lessons with remaining lessons in Table 4 reveals that sizeable differences occur between high and lower scorers on identifiable scales except for "teacher encourages originality."

[4] This is true only with respect to teachers' best lessons as only these were used to identify teachers as high or lower scorers. It remains to be seen whether or not high scorers outscore low scorers across all lessons.

Tables 1 through 3 provide mean and standard deviation comparisons between high and lower scorers on each scale across the three different lesson sets, as well as the "degree of difference" that exists between the groups. The degree of difference on any given scale is computed by subtracting the mean of lower scorers from the mean of high scorers and then dividing by the pooled standard deviation. Typically this statistical computation is called the "effect size". Since this study does not involve control and experimental groups to determine the "effect size", this statistic is referred to as the "degree of difference." Using a conservative estimate by experimental standards, differences exceeding 1 standard deviation between the two groups on any given scale will be considered sizeable or meaningful.

As an example, across all lessons the degree of difference between high and lower scorers on the scale "students giving reasons" is 1.28 standard deviations. Assuming the data is normally distributed, this means that if a teacher were to move from the mean score of the low group to the mean of the high group, it would be equivalent to moving from the 50th percentile to the 90th percentile of all the teachers combined. Frequency distributions can also provide a sense of the rather large differences that exist between the two groups of teachers. Again using the dimension "students giving reasons" as our example, 41 of 44 lessons of high scorers or 93% received a score of '3' or greater on students giving reasons whereas only 19 of 44 lessons of lower scorers or 43% received a score of '3' or more (see Table 5). A sizeable or meaningful difference clearly seems to exist between the two groups on students giving reasons, but note that the 1.28 degree of difference is not particularly large when compared to the differences that emerge on other dimensions of thoughtful practice (see Table 1).

Presentation of findings on each dimension will be followed by an example of a lesson receiving a high or outstanding score (i.e., "4" or "5") and an example of a lesson receiving a low score (i.e., "1" or "2"). In cases where the degree of difference is sizeable and in favor of high scorers, the outstanding example will be taken from a high scorers' lesson and the non-exemplary or low scoring example will be taken from the lesson of a lower scorer. On scales where sizeable differences between the two groups failed to emerge, or on scales where sizeable differences were found but both groups scored relatively low (i.e., a group mean of less than '3'), only an outstanding example will be provided--with the obvious inference that lessons receiving a low score failed to manifest the dimension under discussion.

Few Topics Discussed

With respect to all lessons and teachers' best lessons, sizeable differences were found between high and lower scorers on the dimension of "few topics". A high score indicates that few topics were addressed during a class session while a low score indicates that many topics

were covered. The degree of difference between the two groups exceeded 1 standard deviation for both lesson sets.

As an example, one high scorer focused class discussion on the problem of creating material well-being for everyone in society by analyzing the following proposition: "We live in a world of limited resources but are creatures of unlimited desires." An example of a lesson where many topics were discussed can be found in a lesson on the abolitionist movement in the United States. Discussion ranged across such topics as relations between the U.S. and the countries of Iran, South Africa, and Nicaragua, and how the abolitionist, women's rights, and prison reform movements all fit into Jacksonian democracy.

Substantive Coherence of Lesson Content

Large differences appeared between high and lower scorers with respect to all lessons and teachers' best lessons. The degree of difference between the two groups approached 2 standard deviation for both lesson sets.

An example of a high score on substantive coherence can be witnessed in a high scorers' lesson on the nullification debate between South Carolina and the federal government during the Jackson administration. The class first discussed the meaning of the term nullification. Then arguments for nullification as presented by vice-president John C. Calhoun were explored, followed by an analysis of president Jackson's counter arguments. Once the arguments on both sides were clearly understood, students were asked to generate hypothetical situations today where the nullification issue could arise and where a state might be justified in its actions.

A low score on substantive coherence can be observed in a lower scorers' lesson on president Theodore Roosevelt. The lesson began with students generating criteria or qualities of "leadership greatness" to help them assess T.R.'s leadership as president. Students then began to provide evidence from their readings on T.R. relevant to the criteria generated, but suddenly the central question changed from leadership to whether or not T.R. should be considered a "great Progressive". Shortly thereafter the teacher again redirected discussion by asking if being a reformer is innate or can be learned. For the remainder of the period, discussion switched from T.R.'s position on trusts to a discussion of his leadership greatness and then back again to his status as a great Progressive. The bell rang with no attempt at closure on the teacher's part.

Students Given Time Before Responding

Across all lessons both high and lower scorers gave students sufficient time to think before responding, whereas a sizeable differences in favor of high scorers emerged with respect to teachers' best lessons. Examples of exemplary and non-exemplary lessons are not

provided because high and lower scorers do not differ across all lessons on this dimension, and because "giving students time" is self-explanatory.

Teacher Provides Students With Challenging Tasks

Across all lessons high scorers provide students with more challenging tasks than lower scorers, with the degree of difference again approaching 2 standard deviations. However, with respect to teachers' best lessons differences are less pronounced and fail to exceed 1 standard deviation. Apparently, high scorers more consistently provide challenging tasks for students, but the degree of challenge is about the same with respect to teachers' best lessons.

An example of a highly challenging task can be seen when a high scorer had students assess whether or not Andrew Jackson's presidency enhanced or diminished the level of democracy in the United States. Using criteria of democracy the class had developed in a previous lesson, students had to identify, describe, analyze, and assess various aspects of the Jacksonian presidency (e.g., the election, the inauguration, the spoils system, his view of the presidency, etc.) for its democratizing or non-democratizing effect on the U.S.

A low score on challenge can be observed in a lower scorer's lecture-recitation class session. Students were occasionally asked recall type questions to indicate their retention of material presented in the text or by the teacher. Many terms were defined such as impressment, intervention, Barbary pirates, Leopard Incident, etc. The teacher also outlined for students the views of past presidents on the issue of isolationism.

Teacher Models Thoughtfulness

Across all lessons and best lessons high scorers model thoughtfulness more than lower scorers. Differences exceed 1 standard deviation with respect to both lesson sets.

An example of a high scorer modeling thoughtfulness in an exemplary fashion can be witnessed in a lesson on the rise of anti-Semitism in Germany between 1918 (Treaty of Versailles) and 1941 (the beginning of the Holocaust). The teacher modeled thoughtfulness in the following ways:

-complimented students at the beginning of the period on their thinking during yesterday's discussion

-acknowledged that the textbook presents many details without a unifying structure, making it difficult to determine what information is important

-insisted that students give explanations and reasons for their comments, and where appropriate would direct the class to the section of the reading where the evidence could be found

-encouraged students to recall information and ideas without their notebook, but if unsuccessful, they could then refer to their notes

-acknowledged the difficulty of defining the term "statistic" and of appreciating the magnitude of such numbers as 1 or 10 million lives

-complimented the reasoning sequence underlying a student's answer even if the answer was incorrect (e.g., because the student did not possess some critical piece of information)

-complimented students when good questions were raised

-emphasized ownership of ideas and questions by students (e.g., "Jeff's question" or "Sue's position")

-exhibited certain mannerisms associated with thoughtfulness (e.g., eyes looking toward the ceiling in an act of reflection, head cocked at an angle to indicate a more careful listening posture, stroking of the chin while deliberating, enthusiastically nodding his head while student speaks, etc.)

Much less modeling of thoughtfulness can be observed in a lower scorers' lesson on Frederick Jackson Turner and his "Frontier Thesis". The aim of the lesson was to have students consider the following question: "How has the Frontier Thesis influenced the development of America?" Teacher modeling was deficient in the following ways:

-the central question itself was vague and open to various meanings and interpretations (e.g., how can a thesis developed long after an historical period influence the development of that period?) -class discussion strayed from the central question without direction or purpose

-the teacher did not attempt to connect student comments back to the central question

-the teacher did not share with the class his thinking about any of the points or issues raised by students, or about the central question

-the teacher never sustained his devil's advocate role beyond an exchange or two, resulting in a schizophrenic advocacy of competing and at times contradictory positions that precluded an analysis of statements and positions

-the teacher tended to focus on students' conclusions rather than on the reasoning for their conclusions

-there was no closure or summary of the ideas discussed during the lesson

-the lesson was essentially a "show and tell" of ideas on the part of students, with no attempt on the teacher's part to orchestrate a discussion exhibiting movement and purpose

Students Give Reasons and Explanations for Statements

Across all lessons the students of high scorers offer reasons more consistently than do the students of lower scorers. With respect to teachers' best lessons, however, the means of the two groups are very similar and the degree of difference correspondingly small. Recalling that 3 of 5 lower scorers' classes were composed primarily of high achieving students compared to only 1 of 5 among high scorers, it is not surprising that on certain lessons 'reason giving' by lower scorers' students is comparable to that of high scorers' students. Overall, though, high scorers more consistently require students to provide reasoning for statements.

An example of a lesson where students consistently supported their statements with explanations and reasons is found in a high scorers' economics class. A student responded to the teacher's yes/no question with a response of "no". Though the answer itself was correct, the teacher did not comment. Instead he raised his right arm and rotated his index finger in a circular fashion. Instantly the student blurted out "because!", and after momentary pause for reflection he proceeded to give reasons for his "no" response. This incident illustrates the extent to which students had been trained to provide supporting evidence for their statements, as a simple hand gesture by the teacher could now prompt the desired behavior.

Lack of support for statements can be observed in a lower scorer's lecture-recitation lesson. The teacher reviewed for students 9 social problems attacked by Progressives (e.g., child labor, factory safety, education, etc.), as well as the solutions Progressives proposed and the extent of their success. The students were required to follow along with a worksheet and write down each of the problems, the proposed solutions advocated by Progressives, and the relative success of the proposals as articulated by the teacher.

HOTAV -- Minimal Requirements of a Thoughtful Lesson

Because of the rather large differences observed between the two groups on most of the above scales with respect to both sets of lessons, it is not surprising that sizeable differences, again in favor of high scorers, emerge between the two groups on HOTAV.

We now turn to the remaining four dimensions of thoughtful discourse not used to select teachers as either high or lower scorers.

Careful Consideration of Reasons and Explanations

Across all lessons a large difference in means emerged between high and lower scorers in favor of high scorers, with the degree of difference approaching 2 standard deviations. Similar findings emerged for teachers' best lessons.

An example of a high score on careful consideration can be witnessed in a discussion of the factors that led to a large U.S. national debt in a high scorers' economics class composed of a diverse range of students. Each potential factor was carefully analyzed to see if it could indeed help cause the national debt. An example exchange between the teacher and the class will highlight the dimension of careful consideration of ideas and reasons. Having agreed as a class that an increase in population could contribute to an increase in debt, the class begins to consider "increase in average age" as a potential factor:

Teacher: Why would an increase in average age cost more?

Student: Because more people are staying around more. Because there are more people and so you stay around longer. Then there are more people and that way more people get social security and other things and you have to pay for these people because they're still alive.

T: O.K. So it's not only that there are more of them, but they're different. How are they different?

S: Well, they live longer.

T: Alright, they live longer, so?

S: So they collect more money.

T: Why should they get more money than younger people? Because they're old and wrinkled and have white hair (pointing to his own)?

S: Because they worked already.

T: They worked already. So now what? [Pause--no student responds] Someone's got to help us out here. She says there are more of them. But we already know that because there are 240 million today rather than 76 million in 1900. But it's only half the story that the population got bigger. But it also got older.

S: What was the average age back then and now?

T: Life expectancy then was roughly 56 years; and now for the entire population around 73 years. So, yes, there are more older people, the average age of the country is a lot older, especially in the

last two decades. [Pause] So, what does this mean in terms of our overall question? Why are the government's expenditures going up? How does the fact that they're getting old drive up expenditures?

S: The government has to pay for that.

T: For what?

S: Their house, their

T: They've already paid off their home mortgage. They're retired.

S: Their medical, and they go to the hospital so many times.

T: O.K., that's two. See you knew all the time. Alright, what else?

S: The average retirement age hasn't changed and they'll have a longer time to receive social security benefits, also there are more people of a working age who will receive benefits.

T: O.K. Let's put this on your list of factors as, 'Population grows and ages.'

Failure to engage students in careful consideration can be witnessed in a lower scorers' lesson with honors students. The topic for the day was political leadership, and students were to individually select what they believed were the 3 most important leadership traits from a list of 5 on the blackboard (e.g., charisma, moral integrity, political savvy, etc.) and to add one of their own. Discussion began with students stating the trait they added, and then a lengthy hand count tally was taken to determine which 3 of 5 traits received the most votes. Because students were not required nor given the opportunity to provide reasons for their selections, careful consideration of reasons was absent. Following the hand count the teacher for the remainder of the lesson overviewed with students an upcoming assignment.

Teacher Engaged Individual Students In Socratic Dialogue

Though neither group consistently scored high on this dimension, sizeable differences emerged between the two groups in favor of high scorers across all lessons and teachers' best lessons. The degree of difference exceeded 1 standard deviation for both lesson sets. The criterion for a Socratic exchange was 6 uninterrupted exchanges, 3 each, between teacher and a student (or between two students). Only an example of a lesson scoring high on this dimension will be provided because it would not be instructive to summarize a class where such exchanges failed to occur.

An example Socratic dialogue occurred in a high scorers' U.S. History class comprised of low achieving students. Students assumed

the role of U.S. senators on a Congressional committee interrogating Andrew Carnegie (role played by the teacher) about his monopolistic business practices. One exchange between Carnegie and student-Senator Morris was as follows:

Carnegie: Senators I am a very, very rich man. What do I own?

Morris: You own nearly everything...you own steel, cement, machinery, railroads, and mills.

Carnegie: I own nearly everything but not everything. But I have worked hard for it. I have earned it. Don't tell me I am wrong to own it. I worked my way up from the poor little kid destitute on the streets...and now I am proud of what I won. I own 75% of the steel industry why shouldn't I have 75% of the market? Why shouldn't I have 75% of the territory?

Morris: That's not fair. That's unscrupulous.

Carnegie: What do you mean unscrupulous Senator. Haven't I earned everything I own?

Morris: You would be able to eventually own everything and nobody would be able to have a business.

Carnegie: But what do you mean? If I could own all of industry why shouldn't I? Isn't this a free country? Isn't competition what it is all about? Don't I pay my workers and set fair prices?

Morris: But nobody could compete with you and I don't think there would be any business after that for anybody but you.

Carnegie: Senator you are a difficult man. You are really a difficult man...

Teacher Invites Original Responses

The mean scores for both groups on both sets of lessons are low. Only 13 of 44 high scorers' lessons and 5 of 44 lower scorers' lessons received a score of "3" or greater on this dimension of thoughtfulness. The degree of difference was not sizeable on either lesson set.

One of the 3 lessons in the entire sample of 88 lessons to score a "5" occurred in a high scorer's freshman social studies course. The intent of the lesson was to show students how a society's values influence its political, economic, and social arrangements. The lesson began with a teacher led discussion on what it means to say every culture has certain values. To highlight this idea students were asked to consider how two political systems would differ if one valued equality and the other did not. Though this task invited original

responses from students, the task that followed required them. In small group discussion each group was to construct and describe the workings of a subsystem or aspect of an imaginary, ideal society (e.g., its educational, medical, tax, system). A set of questions was provided by the teacher to help direct each group's discussion (e.g., Would everyone be taxed the same amount? What would be taxed? Would individuals and corporations be taxed at the same rate? etc.).

Teacher Reveals That Disagreement Exists Between Authoritative Sources

Here again both groups scored consistently low with respect to both lesson sets, but with respect to teachers' best lessons, the mean of high scorers was 1.46 vs 2.57 for lower scorers. For the first and only time a sizable degree of difference in favor of lower scorers emerges; i.e., 1.56 standard deviations. Because neither group scored particularly high on this dimension only an example of a lesson receiving a high score will be provided. Only 5 lessons received a score of "5", 4 by high scorers and 1 lesson by lower scorers. The following example is taken from a high scorer's lesson.

In a psychology lesson, students critiqued Carl Rodgers' "client centered" therapeutic approach from both a Freudian and a Jungian "therapist centered" perspective. Students were made aware that experts within the field of psychology disagree fundamentally on treatment methods, including assumptions about the ability of the patient to heal him/herself, about the need to delve into the patient's past, about the need for a forceful, directive therapist, etc.

Summary

What do the above findings suggest with respect to the three questions focusing our analysis? To each of these questions we now turn.

1.) With respect to all lessons and teachers' best lessons are there sizeable differences between high and lower scorers on the first 6 dimensions of practice constituting the minimal requirements of a thoughtful lesson?

Examining teachers' best lessons, we found high scorers' classroom practice to be superior to lower scorers' on; "few topics", "content coherence", "time to think", and "teacher modeling thoughtfulness". Sizeable differences were not found with respect to "students giving reasons" and "teacher providing challenging tasks". Lower scorers, at least with respect to their best lessons, seem to construct equally challenging tasks for students and solicit from students responses that reveal their reasoning. However, compared to high scorers they tend to do much less modeling of thoughtfulness, and cover a greater range of topics during a lesson with much less substantive coherence. More

topics covered may explain why students of lower scorers are given less time to think.

With respect to all lessons, the sizeable differences found above on teachers' best lessons are maintained except for the dimension "time to think". Both groups continue, however, to give students sufficient reflection time before responding (i.e., mean scores of greater than '4'). Substantial differences emerge on two other dimensions of practice; i.e., high scorers give students significantly more challenging tasks and much more frequently require that students provide reasons for their statements. With the addition of challenging tasks and the giving of reasons, across all lessons sizeable differences between the two groups exist on 5 of the 6 dimensions constituting the minimal requirements of thoughtful practice.

2.) With respect to all lessons and teachers' best lessons how do high and lower scorers compare on the remaining four dimensions of thoughtful practice? Are there sizeable differences, and if so, in favor of which group?

Concerning teachers' best lessons, sizeable differences emerged between the two groups of teachers on three of the four scales. High scorers to a much greater extent engage students in "careful consideration of reasons" and "Socratic dialogue", whereas lower scorers are more likely to reveal to students "disagreement among authoritative sources." Recalling that both groups solicit reasons from students with about equal frequency on their best lessons, it appears, however, that high scorers are more likely to go beyond the sharing of reasons and actually exam the merits of those reasons through careful consideration. Careful consideration may explain the substantial differences also found on Socratic dialogue, as Socratic probes often accompany scrutiny of ideas and arguments.

Though both groups scored relatively low on the "authoritative sources" dimension (i.e., less that "3"), how does one explain lower scorers' superior performance on this in their best lessons? One possible explanation involves lower scorers' greater propensity for content coverage and reduced efforts at careful consideration. It may be that the expanded coverage sweep of lower scorers increases the likelihood that students will be exposed, albiet superficially, to differing interpretations and perspectives. Conversely, high scorers' emphasis on depth of coverage and careful consideration with respect to their best lessons may reduce the likelihood that competing perspectives will be discussed in a given lesson.

Across all lessons high scorers again display superior performance on the dimensions of careful consideration and Socratic dialogue. Lower scorers, however, do not maintain superior practice on "authoritative sources," with neither group revealing much of this dimension (i.e., mean scores of less than "2"). As with teachers' best lessons, teacher solicitation of original responses from students

failed to reveal differences between the two groups of teachers across all lessons.

3.) Do high scorers consistently facilitate thoughtful discourse more so than lower scorers; that is, are there sizeable differences between the two groups regardless of lesson set analyzed?

Across all three lesson sets high scorers exhibit superior performance on the same 6 dimensions, and comparable performance on the remaining 4 dimensions.[5] Across all lessons sizeable differences in favor of high scorers emerge on 8 of 10 dimensions. Even if one removes the effect of teachers' best or select lessons on each group's total sample of lessons by only considering teachers' less-thoughtful or "remaining lessons", all of the sizeable differences remain (see Tables 3 and 4).

A final check of stability or consistency across lessons is performed by comparing the mean of teachers' best lessons with the mean of their remaining lessons on each dimension of practice (see Tables 5 and 6). If the degree of difference on a given scale is found to be less than a standard deviation, then substantial differences do not exist and we can conclude that the respective group of teachers were consistent in their practice on the respective dimension. Among high scorers, a sizeable difference emerges on only 1 of 10 scales (i.e., time to think). For lower scorers sizeable differences emerge on only 3 of 10 scales (i.e., teacher modeling thoughtfulness, students giving reasons, and the questioning of authoritative sources). We can therefore conclude that high scorers compared to lower scorers on most practice dimensions consistently display superior performance regardless of the lesson set analyzed (i.e., all, best, or remaining lessons).

COMPARING COMPLETE LESSONS OF HIGH AND LOWER SCORERS ACROSS MULTIPLE DIMENSIONS OF THOUGHTFULNESS:

Having presented quantitative differences between the two groups and provided brief examples of lessons scoring either high or low on each of the 10 dimensions of thoughtful practice, four representative lessons (2 from each group) are now summarized to provide a more concrete, encompassing understanding of the differences in practice that exist between the two groups. A U.S. history lesson of a high scorer is paired with a U.S. history lesson of a lower scorer, and a high scorer's non-history lesson (i.e., psychology) is paired with a lower scorer's non-history lesson (i.e., Civics). Lessons chosen need not be controlled for students' prior achievement because a majority of high scorers' classes contained students of diverse or low prior achievement in social studies whereas lower scorers had a majority of

---

[5] Except on authoritative sources and teachers' best lessons where lower scorers exhibit superior performance.

students with high prior achievement. This finding suggests that being classified as a high scorer did not depend upon working with a more receptive student population. Nonetheless, the paired U.S. history lessons both contain students of high prior achievement. The high scorer non-history lesson contains students of diverse prior achievement while the lower scorer non-history lesson comprises students primarily of low prior achievement.

History Lesson of a High Scorer with Students of High Prior Achievement.

The following lesson is part of a unit on the Constitution in Hilary's 10th grade U.S. History class. One question guides inquiry for the entire unit; i.e., "Were the framers of the Constitution primarily motivated out of self-interest or for the good of the country?" At the end of the unit students must write a paper expressing their position on this question. Students are analyzing information about a sample of 11 framers to generate conclusions about all 55 framers. Today's lesson has three goals: a) review the thinking skill of "organization" and introduce the thinking skill of "synthesis"; b) determine the various political goals of individual framers and make generalizations about the framers as a group; and, c) have students write a thesis sentence (i.e., position statement) for their upcoming paper.

Hilary begins the lesson by reviewing the 4 steps involved in organizing information learned yesterday. The four steps are:

1) Identify your goal (i.e., what will you do with the information?)
2) Survey your data
3) Determine your categorizes of organization based upon your goal
4) Place information in the proper categories

She then asked students if they had any problem with the homework assignment in which they were to complete a chart using the categories of organization they had generated in class the day before (e.g., name, occupation, degree of wealth, how wealth was achieved, political goals, etc.). Students had problems inferring from the readings the political goals of some framers. Rather than provide students with answers, Hilary raises questions and offers factual hints until students make inferences and reach conclusions. For example, students come to see that James Madison feared the self-interest of many of the framers and therefore desired compromise between the various factions in order to stabilize the country and hold it together, or that Edmund Randolph wanted a balance of power between the 3 branches of government but that each branch be run by members of the educated and monied elite. Hilary addressed other questions in a similar manner, specifically students' moral indignation at the thought of greedy conventioners like that of Jonathan Dayton:

Student $_1$: Why did they have Jonathan Dayton there, he was such a jerk!

Student $_2$: Yea.

Hilary: What do you mean he was a jerk?

Student $_1$: He was a crook trying to cheat the government while they're trying to make a good government.

Student $_3$: Do you think it was because he was a war hero or something?

Student $_4$: I'm sure they all cheated that way, why else would they be there?

Hilary: Who was selecting these people to go to this Constitutional Convention?

Student $_6$: The people.

Student $_5$: Themselves!

Hilary: Sure, they were an elite group selected by an elite group. You have to be able to vote in order to vote for somebody to go to the Constitutional Convention. So you're dealing with a select group here. We would love to believe that then and now all people who are involved in politics are honest and just but we know that's not the case..."

Another analysis of a framer's intentions is as follows:

Hilary: Why would Roger Sherman, if he achieved his position in life through his own initiative want as one of his political goals to keep the power in the hands of an elite? What do you think?

Student 1: Maybe because he made it on his own he felt he was better than the rest of the people.

Student $_2$: He was an exception.

Student $_3$: He sort of educated himself. He worked hard to get where he was and didn't think people should get to vote just by not doing anything.

Hilary: Alright. Anything else? Where had he started? What does he think he knows about those people?

Student $_3$: [mumbled-inaudible response]

Hilary: You see what she (i.e., student$_3$) is saying? We don't know his political goals but we can make some assumptions. One assumption Jim is making is that Sherman has already lived with these people who are not concerned with government

affairs and he knows this. Than Sherman's concern is that they not assume such a position (i.e., political decision-making). What would Sherman assume are the major concerns of farmers and day-to-day working class people? What's their major concern?

Student $_4$: Their land

Hilary: Sure, they've got to survive. Having the benefit of time to play politics is a luxury. And maybe he is saying that these people cannot do this. Right? And therefore politics should left in the hand of people who are educated to that kind of position.

Having satisfied students' concerns regarding the political goals of various framers, Hilary introduces the thinking skill of synthesis: "The next step, now that we've organized all of this information, is to synthesize, to put it together to make something new." Various analogies are used to clarify the concept such as the ingredients or parts that go into making a pizza or the whole. The three steps of synthesis are explained:

1) Look at your organized information
2) Decide what is in common
3) Pull the common information together into a generalization

Students are now asked to synthesize or generalize across each category of the homework chart starting with "name/background". Three students contribute and conclude that the framers are men, white, landowners, and voters, with English backgrounds. Generalizations are also made regarding "occupation", "degree of wealth", and "how wealth was gained". The class now arrives at the critical generalization with respect to their upcoming position paper. Hilary asks: "When you look at their political goals, what general features can you find in the majority of delegates?" One student suggests that the framers were out to "benefit the elite", while another student states that they wanted to develop "rules and laws that would benefit themselves." The class seems to be in agreement here, so Hilary asks if there are any exceptions. Discussion leads to the conclusion that Madison, Franklin, and Hamilton are the exceptions.

Students are then given 5 minutes to "take the various common elements" and combine them into "a general statement" which will serve as the thesis for their position paper. Hilary asks for volunteers to share with the class their thesis sentence. Six are read, two of which are as follows;

"The framers of the Constitution were interested in developing regulations that were beneficial to their wants and needs."

"The creators of the Constitution were members of the elite class representing the elite to benefit the elite."

With some students Hilary suggests minor but important modifications. For example one student refers to the framers as "elite delegates" in her thesis sentence. Hilary points out that the phrase "elite delegates" implies that there were also "non-elite delegates" which was not the case. She suggests the student use an alternative phrase such as "the delegates were part of the elite".

Before the class ends 4 additional points are addressed:

-Hilary informs students that generalizations do not always contain arguments, and then asks students to determine if indeed their generalization contains an argument.

-Hilary offers students a way to check for an argument by assessing whether or not there is an alternative position to their own?

-Hilary reminds them that they're looking at a sample of 11 framers from a total of 55, and therefore the evidence they supply to support their generalization must apply to most of the 11 framers or it won't be convincing.

-Hilary explains the homework assignment in which students are to outline the body of their position paper by providing the ideas and facts that will support their thesis.

Reviewing the above lesson with respect to the observational for scales, many outstanding features emerge. Few topics were discussed, specifically the thinking skills of organization and synthesis and their application to the question of intent of 11 Constitutional framers. The content was presented and discussed in a coherent manner, with an eye always focused on the relationship between the data and the overriding unit question. The teacher modeled thoughtfulness when summarizing the inductive steps students used to arrive at the political goals of various framers, when sharing with the class her inferences about framers' political goals, and when giving ownership to ideas expressed by students (e.g., "one assumption Jim is making..."). The problems posed were challenging as evidenced by the degree of inference required. Students' responses were carefully analyzed for factual accuracy, correctness of inference, and clarity of generalization in their thesis statement. Students were given time to think before questioning and responding, and Socratic exchanges occurred on occasion between Hilary and a few students when assessing framers' political goals and students' thesis statements.

U.S. History Lesson of a Lower Scorer with Students of High Prior Achievement

Larry's year-long Advanced Placement United States History class is composed of high achieving juniors. The lesson focuses on the 1848 Seneca Falls Declaration of Women's Rights, and occurs at the end of a unit on the broader theme of Jacksonian democracy. The goal of the lesson is to answer the following question: Did American women have a case for equal rights during the Jacksonian period?

At the start of class students received a one page handout with excerpts from the 1848 Seneca Falls Declaration and the 1776 Declaration of Independence, and a reproduction of a partial chapter (i.e., 6 pages) on the rights of women from a U.S. history college textbook containing relevant quotes by Abigail Adams and her son. A teacher-centered class discussion began once students finished reading the page of excerpts from the Declarations of 1776 and 1848.

Discussion opens with Larry asking the class to identify the ways in which the two documents can be compared. Students eventually agree to have the following summary statement written on the blackboard regarding the Declaration of Independence; "People have the right to alter or destroy the government". With respect to the Seneca Falls Declaration, students agree to the summary statement that "Women have the right to refuse allegiance to a government." Larry asked the class if these two statements meant the same thing, but the question was not pursued. Instead the class moved on to a second point of comparison; students paired "All men are created equal" from the Declaration of Independence with "All men and women are created equal" from the 1848 Seneca Falls Declaration.

At this point a student, Peter, suggests that Jefferson and the other revolutionaries would have been laughed at if they included women in the document, especially since they were directing their message to British leaders who were all men. He adds that the Seneca Falls Declaration occurred once "we were our own country, and the people of our country were saying men and women are equal." Larry immediately asks the class if they agree with Peter's statement. Two students offer brief comments that seem to disagree with Peter (inaudible on tape). Larry then jokingly informs the class that "...if I were alive at the time you know for sure I wouldn't have backed women's rights." Two female students object and call him chauvinistic. Larry qualifies his comment by stating, "I said in 1848. Now it's 1986." He then redirects the class by requesting other points of comparison between the two documents as "a foundation" is needed to help them answer the central question of the day's lesson.

A student mentions that both documents allude to a "history of repeated injuries." Specific examples of injuries are cited such as taxation without representation with respect to the Declaration of Independence, and no vote, no law making power, and no property rights with respect to women and the Seneca Falls document. Discourse then proceeds as follows:

Larry: What's the problem with these injuries?

Student: Men would say this is no big deal.

Larry: What else is the problem, what is the problem with this document?

Student: These injuries are still going on today.

Larry: Maybe, but what else?

Student: They had no power to enforce it.

Larry: O.K., what else?

Student: Nobody would pay attention to it.

Larry: Who wrote the document?

Student: Women.

Student: But men wrote the Declaration of Independence.

Larry: Oh, there are many men today who come on to the women's side. Not me, but there are men who would. But do you find that this might have been one-sided this time--very one-sided? In other words women didn't have a case for equal rights.

Larry then asks the class what they think. Two students offer comments, but neither is a direct response to Larry's statement. He then abruptly changes the direction of the discussion, either asking two different questions in one sentence or modifying the initial question in the latter half of the sentence: "Which right do you think was violated the most, which affected women probably the most?" One student cites lack of participation in law making. A second student mentions lack of higher education, and how this would result in women not "knowing how to present a good argument. Larry immediately asks her if that is a social or economic argument to which the student responds "both." He then calls on one more student to share her thoughts before asking the class to turn to their second handout.

A student reads aloud a statement supporting women's rights by the son of President John Adams and his wife Abigail to members of the House of Representatives in 1838. The quote ends with Adams posing the following questions: "Are women to have no opinions or actions on subjects leading to the general welfare? Where do the gentlemen get this principle?" Larry asks students "what do you think about that?" One male student, Jim, asserts that once women starting working, unemployment increased. A second male student agrees with Jim that "society ran more smoothly but in the long run having women work benefits society." At this point Carol asks, "Who gave men the right

to take the power?" Larry ignores her question and asks, "Is there any man here who would stay home if his wife could make a better living?" Two male students respond, one suggesting he'd be bored and the other stating he'd quit if he was a secretary or garbage collector.

Larry then reintroduces his central question of the day, "What do you think was the most critical case women might have had for equality?" A student misinterprets his question, suggesting that the right to vote is most important because "it gives you the opportunity to change other things." Rather than rephrase the question Larry moves on to another question, "Is there anything in the present time that justifies women in terms of their case for equality?" A few female students point out that women possess equal brain power, and that women should receive the same pay as a man if doing the same job. A third female, Sue, asserts that women are inferior physically and that it's wrong to lower the standards for policemen.

Larry then proposes to the class that "we try to compose a possible amendment that you would put in regarding women's rights." The following discourse ensues:

Student: I don't think it's right to have an amendment. You can't expect this to happen over night. It has to be a gradual evolution.

Larry: Why should women have to wait?

Student: An amendment is not going to change anything.

Larry: Why don't the women revolt then?

Student: If we start giving the women these rights then we have to give everyone else these rights.

Larry: How many people are in favor of the Equal Rights Amendment?
[hand count reveals some females and no males favor it]
Why would a woman not be in favor of ERA?

Various reasons are then offered by Larry and a few students (including females), such as women may not want to be drafted, their husbands may be entitled to alimony in a divorce, etc. The bell rings.

Reviewing the above lesson with respect to the 10 dimensions of thoughtfulness, very few positive features emerge. Few topics were discussed, specifically the issue of women's rights in 1848 and today. However, there was very little content coherence as discussion jumped from a comparison of Declarations, to assessing the merits of the Seneca Falls argument, to determining which right was most important for women, to the reading of a quote, back to the central question,

and finishing with a series of questions regarding women's rights today and the construction of a new ERA amendment. The central question was posed on three occasions but within minutes it was lost to digression. Larry failed to model thoughtfulness, because he did not: a) share his reasoning on the various issues addressed--except on the central question where he dismissed the Seneca Falls document on the questionable grounds that only women advanced it; i.e., the document was "very one-sided; b) compliment or encourage students' on their thinking; c) respond to seriously some of the questions raised or comments made by students; or, d) require that the student respond to a preceding comment or question. Students usually did provide their reasoning, in part because it is a department policy and because they are high achieving students. Not once, however, did Larry request that a student provide additional support for a statement or add further embellishment. Because there was nothing more than a "show and tell" of statements and reasons, careful consideration did not take place. Since Socratic exchange is often involved in careful consideration it is not surprising that it was not also observed. Some of the questions posed and issues raised were quite challenging, but because students were given little time to think before responding and because there was no careful consideration, students were not made aware of the difficulty of some questions. Original responses were not solicited by the teacher and students were not shown that disagreement exists among authoritative sources.

Psychology Lesson of A High Scorer with Students of Diverse Prior Achievement

This psychology class lesson is part of a unit on behavioral psychology. Hans has two goals for today's lesson: a) enhance student understanding of classical conditioning and its principles (i.e., generalization, discrimination, and extinction); and, b) highlight the general point or rule about thinking that one must always check and question one's assumptions. Both goals are pursued by placing students in the challenging role of behavioral therapist/problem solver.

Hans begins with a "RAP" (review and preview) of classical conditioning. Students are told of a relationship between dog, Bowser, and his master, Alphonse. "Whenever Alphonse rips off some paper towels and puts on his coat, Bowser knows it is walk time and barks with joy. When Alphonse pulls off paper towels for dinner napkins, however, Bowser does nothing." Students in small groups are to solve 3 problems using the principles of discrimination and extinction: diagram the original classical conditioning that occurred with the leash, explain how the principle of generalization led Bowser to bark with joy when observing the coat and paper towel, and explain why Bowser doesn't bark with joy when Alphonse pulls off a paper towel for use as a dinner napkin.

Following 10 minutes of discussion in small groups the class reconvenes. Little time is spent on the first problem, because few

students had difficulty understanding and diagraming Bowser's conditioned response of joy to the conditioned stimulus of the leash. It is diagramed on the board as follows:

unconditioned stimulus (UCS) ----------> unconditioned response (UCR)
(walking outside) (joy)

conditioned stimulus (CS) ----------> conditioned response (CR)
(seeing Alphonse with leash) (joy)

The second problem proves difficult for more students. One student correctly explains how Bowser, through the principle of generalization, connected the leash to Alphones's coat and the ripping of paper towel. When a few puzzled looks persisted, Hans provided clarification through questioning rather than direct explanation:

Hans: "Without generalization when would the dog bark with joy?"

Student: "When the leash is on his neck."

Hans: "That's right. As soon as the leash is on his neck. You could even have it two inches in front of him and without generalization what would he do, Eddie?"

Eddie: "Nothing."

Hans: "Right. So generalization is the essence of all of this learning."

Following a few additional points of embellishment, Hans directs the class to the third problem ("the hard part"), and requests a "good bull's-eye explanation." The first student-volunteer is told that his answer is "on target but not a bull's-eye." A few other students offer "on target" explanations rather than bull's-eyes, so Hans again intervenes with questions to elucidate understanding. Students come to see that through the principle of discrimination Bowser experiences joy only when the paper towel is paired with the coat. At this point a student asks; "What if the weather changes and you don't need a coat?" Hans rewards her with an enthusiastic "Ahhhh", and states that he too had thought about that possibility. Again, instead of providing an answer Hans asks her what she thinks would happen. When she states "relearn everything", Hans and the class laugh at the thought of how confused the dog must be at certain times of the year.

Having completed the RAP, Hans states that for the remainder of the period, "We're going to have some real interesting stuff go on here. It's going to be fun." Using the blackboard to outline key points from a reading assignment of the night before, he quickly reviews the 3 types of behavioral therapies discussed in the article: forming new conditioned responses, using aversive conditioning to remove negative behaviors, and applying reciprocal inhibition to

remove phobias. He then announces that today the class would focus on the first of the three behavioral therapies (i.e., forming a new conditioned response), specifically how the student-therapists will help Hans stop his bedwetting. Before doing this, Hans quickly reviewed Pavlov's famous classical conditioning experiment with a diagram (i.e., UCS/food ---> UCR/salivation generalizes to CS/bell ---> CR/salivation). [Hans indicated later that this momentary return to Pavlov was designed as a "trap" to ensnare students in their thinking about the upcoming bedwetting problem. The thinking trap was that students would now tend to view the bell as only a conditioned stimulus. However, as will be seen, in order to correctly diagram the bedwetting problem students must define the bell as an unconditioned stimulus rather than a conditioned stimulus as in the Pavlov diagram. Hans then uses this thinking episode to highlight the general point about the importance of questioning one's assumptions when thinking.]

Returning to the bedwetting problem, students are told that Hans' mattress contains a bell that rings whenever the mattress becomes wet. In small groups students are to diagram a conditioning schedule that will help Hans with his problem using some of the following terms: bell, full bladder, urination, and wake-up. Following lively small group discussions, representatives from two different groups go to the board to diagram their group's prescribed conditioning schedule. During discussion various students and Hans find problems with both schedules. For example, one group suggested the following:

UCS ----------> UCR			CS ----------> CR	
(full bladder)	(urinate)	leads to	(bell)	(urinate)

During the analysis Hans laughingly points out that he paid the group $500 for a therapeutic intervention that now has him urinating every time he hears a bell! "You've now created a far bigger problem for me."

Four more groups write their conditioning recommendation on the board, but again class discussion reveals flaws in each. As each attempt fails, Hans provides encouragement to keep students thinking as well as comic relief: "Gee, I hope we can get this or I'm really in trouble...I like to see you guys think. It's a tough one. There is a solution folks, and it's a good one." Finally, Hans must offer a hint. He tells the class that he purposely reviewed Pavlov before they got started on this problem to put them in a mindset that would make solving the problem more difficult. If he hadn't done this, he claims 50% of the students would have solved it already. Within a minute a student provides the correct conditioning relationship, explaining how generalization will lead Hans to associate the bell with a full bladder and waking up.

UCS ----------> UCR			CS ----------->CR	
(bell)	(wake up)	leads to	(full bladder)	(wake up)

Following a few points of clarification, Hans concludes by explaining to students the reason for the Pavlov review and his desire to trap students in their thinking. He states:

"This is a tough one. Some good thinking went on here. I was particularly excited when I saw Shane using what level of thinking? He started eliminating these [i.e., the terms] by saying there are no unconditioned stimulus and responses left. But there was one left wasn't there, bell vs wake up. Shane were you kind of caught in the Pavlov trap? See why I set that up? Just a little point about thinking here. What's one thinking error that we often run into? We don't question our assumptions, right? What was our assumption there? That bells are always a conditioned stimulus...",

The lesson ends with Hans previewing the reciprocal inhibition homework assignment for tomorrow, one based on a Gary Larson cartoon in which students are to make various evaluations of "Professor Gallagher's controversial technique of simultaneously confronting the fear of heights, snakes, and the dark." (See worksheet below). As students are scrambling out the door, Hans makes a final attempt to instill motivation and a good effort; "It's going to be tough. It's a good one. Do a good job on it. I want good work...Good thinking today, Shane..."

Reviewing the above lesson with respect to the observational scales, many outstanding features emerge. Few topics were discussed, specifically classical conditioning with respect to Bowser and to Hans' bedwedding. The content was presented and discussed in a very coherent manner, and the teacher modeled thoughtfulness in his own thinking and in his appreciation and enthusiasm for the thinking of students. The problems posed were challenging, and students' responses were carefully analyzed by the teacher and students. Students were given time to think during small group, and Socratic exchanges occurred on occasion between Hans and students. The lesson did not reveal to students disagreement among authoritative sources, though later in the week students critiqued behavioral therapies from previously studied Freudian and Jungian perspectives.

Civics Lesson of a Lower Scorer with Students of Low Prior Achievement

This lesson focused on the system of checks and balances between the executive, judicial, and legislative branches of the United States federal government. Lloyd began by reviewing some of the points covered in the previous day's class discussion. He asked why the Constitutional Convention wanted to divide power. A student responded, "The government might destroy freedom." Lloyd acknowledged the comment as correct and reiterated the point; "Good, there was a fear they would not have enough freedom." Lloyd then asked students to name the document of 1776, and to name the first plan of government. Students

quickly supplied the correct answers; i.e., the Declaration of Independence, and the Articles of Confederation. Lloyd reminded the class that "in 1787 we went to a federalist system", and then asked them to explain this system. Failing to get the desired response, he stated that the 3 branches of the federal government were set up due to "the fear the people had about the national government becoming too powerful." Students supplied the names of the three branches on Lloyd's request and their respective function (i.e., legislative--"make laws", executive--"enforce laws", and judicial--"interpret laws"). With the review of yesterday's discussion completed, Lloyd told them this topic is important as it would appear on the final exam much later in the school year.

Before discussing the system of checks and balances, Lloyd in a few sentences summarized how the Articles of Confederation proved inadequate, and how the framers wanted a stronger central government but one that could be controlled. He then asked students to open their textbook to a paragraph that summarized how the three branches of government checked and balanced the power of each other. A student was asked to begin reading the paragraph; "The executive branch checks the legislative branch by signing or vetoing a bill." Lloyd then interrupts, asking what the word veto means. Once this word was clarified the student resumed reading to the class, explaining how Congress can override the president's veto with a 2/3rds vote. Lloyd at this point asked students to take time to write in their notes "why we have a system of checks and balances." While students wrote Lloyd read aloud the text's answer to this question; "The purpose of the checks and balances is to make sure that no branch of this national government can become too powerful."

Discussion resumed with Lloyd asking students how the president checks the Supreme Court. A student correctly responded that the president appoints members to the Supreme Court. Without clarifying how the president's power of appointment in fact checks the power of the Supreme Court, Lloyd directed students to the appropriate line of their worksheet where they were to write down "President appoints members of the Supreme Court." Lloyd then asked students how often the Supreme Court checks the President. Following a few incorrect responses by students, he explained that the Supreme Court can tell the president to enforce certain laws and tell him which laws are constitutional and unconstitutional. Lloyd reminded students that the president cannot just go out and make a law; "Congress makes the law. The Supreme Court interprets the meaning of the law and whether it's constitutional or unconstitutional. He has checks. He's not an emperor or dictator."

At this point students are asked to finish the worksheet on the system of checks and balances at their desk. Lloyd informs them he will walk around and lend assistance. 10 minutes later students had completed the worksheet and Lloyd resumed discussion. Students were asked how Congress checks the president. A student correctly

responded, "It can pass laws over the president's veto." Two more questions are asked (i.e., How does the Supreme Court check Congress? and, How does Congress check the Supreme Court), with students providing the desired answer before the bell rings ending class.

Reviewing the above lesson with respect to the 10 dimensions of thoughtfulness, a few positive features emerge. Few topics were discussed, specifically the separation of power among the three branches of government and the system of checks and balances. Although the questions posed required only simple recall, students were given sufficient time to think before responding. Finally, the content, though covered very superficially, was for the most part coherent.

The lesson failed, however, with respect to the other dimensions of thoughtfulness. The tasks posed by the teacher were not challenging, rather students recited from the textbook responses desired by Lloyd. This, in part, led to little modeling of thoughtfulness, as few situations arose where the teacher had to think through a question or problem with students. Lloyd did not encourage or reward students' thinking, nor did he exhibit much enthusiasm or appreciation for the topic and ideas being discussed. Students did not offer reasons for their answers, but instead only supplied phrases from the textbook or the previous day's discussion. Without challenging tasks or reason giving, it is not surprising that careful consideration or Socratic exchanges failed to take place. Finally, original responses were not solicited by the teacher, and students were not shown that disagreement exists among authoritative sources.

COMPARING HIGH AND LOWER SCORERS ON INSTRUCTIONAL FORMAT, READING MATERIALS, AND WRITING TASKS:

As a final point of comparison the frequencies with which high and lower scorers engage students in various instructional formats, and types of reading materials and writing tasks are presented.

Beginning with instructional format, part of virtually every high scorer's lesson contains a whole group teacher centered discussion, compared to 2 of every 3 lessons among lower scorers. Roughly 1 in 4 lessons of high scorers involve students in small group work whereas small group work is virtually non-existent in the lessons of lower scorers. The reverse situation is found with respect to lecture; that is, 1 in 4 lessons among lower scorers include lecture compared to virtually no lecturing among high scorers. Similar findings emerge with respect to recitation, as 1 in 4 lessons exhibited a recitation format among lower scorers compared to no recitation among high scorers. In neither the classes of high or lower scoring teachers was there much student centered discussion, viewing of films or filmstrips, seatwork, or oral reports (see Table 8).

Turning to type(s) of reading materials used during lessons, on average, lower scorers' use a textbook(s) 1 of every 2 lessons, compared to approximately 1 in 5 lessons among high scorers (see Table 9). One in 3 lessons finds both groups using primary source materials, but rarely does either group use articles, literature such as novels, poems, etc. One in 3 lessons finds high scorers compared to 1 in 6 lessons among lower scorers using "other" reading materials. For the most part these are reading materials the teacher has developed using a photocopier or mimeograph machine. Finally, high scorers do not use any reading materials in about one-fourth of their lessons, whereas lower scorers do not use any readings in about one-tenth of their lessons. This may be due to high scorers' greater emphasis on teacher-centered whole group and student-centered small group discussion, combined with lower scorers' greater emphasis on recitation and seatwork activities.

For the most part both groups require some form of writing during class.[6] Neither group is likely to assign paragraphs or essays, research papers, or "other" types of writing tasks (see Table 10). The most frequently assigned type of writing among both high and lower scorers is that of notetaking. Three of 4 high scorers' lessons require notetaking compared to 1 in 2 lessons among lower scorers. High scorers are more likely than lower scorers to require writing complete sentences, but both groups require this infrequently. Worksheet or short answer writing tasks are equally stressed, about 1 in 3 lessons for both groups.

The instructional formats, classroom reading materials, and classroom writing tasks of high vs lower scorers can be summarized as follows. High scorers' lessons almost always include teacher-centered whole group discussion, with an occasional small group activity. One is likely to find high scorers using primary sources and teacher-generated reading materials rather than textbooks. Students of high scorers are likely to be found taking notes, and occasionally completing worksheets and other short answer writing tasks. Lower scorers' lessons are also likely to include teacher-centered whole group discussion. However, unlike high scorers, one is also likely to observe during these lessons a lecture or recitation instructional format. Textbooks are the main reading staple, but like high scorers primary sources materials are also frequently used (34%). Notetaking is slightly less emphasized by lower scorers but common to a majority of lessons, while worksheet/short answer writing assignments are required less often but at a rate comparable to high scorers.

CONCLUSION

---

[6] In only 5% of high scorers' classes and 16% of lower scorers' classes was no form of writing required.

High scorers compared to lower scorers consistently display superior performance on virtually every dimension of thoughtful practice. Differences are slightly less pronounced when teachers' best lessons are compared, yet even here sizeable differences emerge on 6 of 10 dimensions. Across all lessons sizeable differences emerge on 7 of 10 dimensions, 6 of which exceed 1.5 standard deviations (i.e., substantive coherence, Socratic dialogue, careful consideration, few topics, challenging tasks, and teacher modeling thoughtfulness).

Differences between the two groups emerge on other elements of practice. High scorers engage students in teacher-centered whole group discussion almost every lesson compared to only a majority of lessons for lower scorers. Lecture and recitation occur in approximately one-fourth of lower scorers' lessons whereas these two instructional formats are generally not found in the lessons of high scorers. As might be expected, small group discussion occurs more frequently in high scorers' lessons (one-fourth of the time), but lower scorers rarely employ this format. The many sizeable differences found between high and lower scorers on the 10 dimensions of practice may, in part, be attributed to the use of lecture and recitation formats which are less conducive to facilitating thoughtful discourse than whole group or small group discussion.

In light of the disparaging assessment textbooks have received concerning the facilitation of students' thinking, it is interesting to note that lower scorers use textbooks in about one-half of their lessons while high scorers use textbooks in less than one-fifth of their lessons. High scorers on the other hand are twice as likely to provide students with teacher generated reading materials. This may explain, in part, less content coverage (i.e., few topics) by high scorers as the reading materials used are generally less survey oriented than standard textbooks.

TABLE 1

CLASSROOM PRACTICE COMPARISONS ACROSS ALL LESSONS (n=88)

SCALE	(n=44) High Scorers Mean$_1$	S.D.	(n=44) Lower Scorers Mean$_2$	S.D.	$\frac{\bar{x}_1 - \bar{x}_2}{\text{S.D. pooled}}$ Degree of Difference
-few topics	4.55	(.73)	3.59	(1.04)	1.74*
-substantive coherence	4.30	(.73)	2.66	(.75)	1.84*
-time to think	4.23	(.64)	4.20	(.70)	.06
-challenging tasks	4.57	(.70)	3.36	(1.12)	1.91*
-teacher models thoughtfulness	4.55	(.63)	2.80	(.93)	1.84*
-students give reasons	3.93	(.79)	2.82	(1.26)	1.28*
-HOTAV (above 6 scales)	4.35	(.45)	3.24	(.58)	1.57*
-careful consideration	4.09	(1.01)	2.34	(.94)	1.88*
-socratic dialogue	2.84	(1.24)	1.52	(.88)	1.69*
-teacher invites original responses	2.18	(.95)	1.68	(1.01)	.81*
-disagreement among authoritative sources	1.95	(1.33)	1.70	(1.00)	.35

* denotes a statistically significant difference between the two groups at the .05 level.

TABLE 2

CLASSROOM PRACTICE COMPARISONS ON SELECT (BEST) LESSONS (n=20)

SCALE	(n=13) High Scorers Mean	S.D.	(n=7) Lower Scorers Mean	S.D.	Degree of Difference
-few topics	4.92	(.28)	4.29	(.95)	1.16*
-substantive coherence	4.69	(.48)	3.00	(.00)	1.90*
-time to think	5.00	(.00)	4.29	(.49)	1.83*
-challenging tasks	4.77	(.44)	4.43	(.58)	.54
-teacher models thoughtfulness	4.85	(.38)	3.86	(.69)	1.04*
-students give reasons	4.38	(.51)	4.14	(.69)	.28
-HOTAV (above 6 scales)	4.77	(.16)	4.00	(.24)	1.08*
-careful consideration	4.69	(.48)	3.00	(.58)	1.82*
-socratic dialogue	3.38	(1.39)	2.00	(1.16)	1.78*
-teacher invites original responses	1.92	(.64)	2.29	(.95)	-.58
-disagreement among authoritative sources	1.46	(1.13)	2.57	(1.62)	-1.56*

* denotes a statistically significant difference between the two groups at the .05 level.

TABLE 3

CLASSROOM PRACTICE COMPARISONS ON NON-SELECT (REMAINING) LESSONS (n=68)

SCALE	(n=31) High Scorers Mean	S.D.	(n=37) Lower Scorers Mean	S.D.	Degree of Difference
-few topics	4.39	(.80)	3.46	(1.02)	1.69*
-substantive coherence	4.13	(.76)	2.59	(.80)	1.72*
-time to think	3.90	(.47)	4.19	(.74)	-.73*
-challenging tasks	4.48	(.77)	3.16	(1.09)	2.10*
-teacher models thoughtfulness	4.42	(.67)	2.59	(.83)	1.92*
-students give reasons	3.74	(.82)	2.57	(1.19)	1.35*
-HOTAV (above 6 scales)	4.18	(.41)	3.09	(.51)	1.53*
-careful consideration	3.84	(1.07)	2.22	(.95)	1.74*
-socratic dialogue	2.61	(1.12)	1.43	(.80)	1.51*
-teacher invites original responses	2.29	(1.04)	1.57	(.99)	1.17*
-disagreement among authoritative sources	2.16	(1.37)	1.54	(.77)	.87*

* denotes a statistically significant difference between the two groups at the .05 level.

TABLE 4

SUMMARY OF TABLES 1-3*

SCALE	(Table 1) ALL LESSONS	(Table 2) BEST LESSONS	(Table 3) REMAINING LESSONS
-few topics	X	X	X
-substantive coherence	X	X	X
-time to think		X	
-challenging tasks	X		X
-teacher models thoughtfulness	X	X	X
-students give reasons	X		X
-HOTAV	X	X	X
-careful consideration	X	X	X
-socratic dialogue	X	X	X
-teacher invites original responses		X	
-disagreement among authoritative sources		-X	

* An "X" indicates that the degree of difference between high and lower scoring teachers exceeds 1 standard deviation in favor of high scorers, while a "-X" indicates that the degree of difference exceeds 1 standard deviation in favor of lower scorers.

TABLE 5

FREQUENCY DISTRIBUTIONS OF HIGH AND LOWER SCORERS ON THE 10 DIMENSIONS OF THOUGHTFUL CLASSROOM PRACTICE

	High Scorers					Lower Scorers				
	1	2	3	4	5	1	2	3	4	5
-Few Topics	0	1	3	11	29	1	5	15	13	10
-Substantive Coherence	0	0	7	17	20	1	19	18	6	0
-Time to Think	0	0	5	24	15	0	1	4	24	15
-Challenging Tasks	0	1	2	12	29	1	11	11	13	8
-Teacher Models Thoughtfulness	0	0	3	14	27	1	19	14	8	2
-Students Give Reasons	0	2	9	23	10	5	20	1	14	4
-Careful Consideration	1	3	5	17	18	7	20	14	1	2
-Socratic Dialogue	6	13	13	6	6	28	12	2	1	1
-Original Responses	10	21	9	3	1	24	15	2	1	2
-Authorities Disagree	25	6	7	2	4	26	8	8	1	1

TABLE 6

COMPARISON OF HIGH SCORERS' BEST LESSONS WITH THEIR REMAINING LESSONS

SCALE	(n=13) Best Lessons	(n=31) Remaining Lessons	(n=44) S.D.	Degree of Difference
-few topics	4.92	4.39	(.73)	.72
-substantive coherence	4.69	4.13	(.73)	.76
-time to think	5.00	3.90	(.64)	1.71
-challenging tasks	4.77	4.48	(.70)	.42
-teacher models thoughtfulness	4.85	4.42	(.63)	.69
-students give reasons	4.38	3.74	(.79)	.81
HOTAV (the above 6 scales)	4.77	4.18	(.45)	1.32
-careful consideration	4.69	3.84	(1.01)	.84
-socratic dialogue	3.38	2.61	(1.24)	.62
-teacher invites original responses	1.92	2.29	(.95)	-.39
-reveal disagreement among authoritative sources	1.46	2.16	(1.33)	-.53

TABLE 7

COMPARISON OF LOWER SCORERS' BEST LESSONS WITH THEIR REMAINING LESSONS

SCALE	(n=7) Best Lessons	(n=37) Remaining Lessons	(n=44) S.D.	Degree of Difference
-few topics	4.29	3.46	(1.04)	.80
-substantive coherence	3.00	2.59	(.75)	.55
-time to think	4.29	4.19	(.70)	.14
-challenging tasks	3.00	3.16	(1.12)	-.14
-teacher models thoughtfulness	3.86	2.59	(.93)	1.37
-students give reasons	4.14	2.57	(1.26)	1.24
HOTAV (the above 6 scales)	4.00	3.09	(.58)	1.57
-careful consideration	3.00	2.22	(.94)	.83
-socratic dialogue	2.00	1.43	(.88)	.57
-teacher invites original responses	2.29	1.57	(1.01)	.72
-reveal disagreement among authoritative sources	2.57	1.54	(1.00)	1.03

TABLE 8

COMPARISON OF CLASSROOM INSTRUCTIONAL FORMAT (n=88)

Format	High Scorers % of lessons	Lower Scorers % of lessons
-teacher centered discussion	96	66*
-student small groups	23	2*
-student centered discussion	2	0
-lecture	0	27
-film	2	0
-recitation	0	23
-seat work	0	5
-oral reports	2	2

* denotes a statistically significant difference between the two groups at the .05 level.

TABLE 9

COMPARISON OF CLASSROOM READING MATERIALS (n=88)

Materials	High Scorers % of lessons	Lower Scorers % of lessons
-text	18	52*
-primary source	30	34
-article	7	6
-literature	0	2
-other	36	16*
-none	23	9

* denotes a statistically significant difference between the two groups at the .05 level.

TABLE 10

COMPARISON OF CLASSROOM WRITING TASKS (n=88)

Format	High Scorers % of lessons	Lower Scorers % of lessons
-outline	0	0
-class notes	73	55
-worksheets/short answer	30	34
-complete sentences	21	9
-paragraph or essay	0	7
-research paper	0	0
-other	5	0
-none	5	16

* denotes a statistically significant difference between the two groups at the .05 level.

V

STUDENT PERSPECTIVES ON COGNITIVELY CHALLENGING CURRICULUM

Robert B. Stevenson

INTRODUCTION:

Much research in high schools indicates that students are rarely challenged to use their minds (Cusick, 1973; Goodlad, 1984; Powell, Farrar, & Cohen, 1985; Sizer, 1984). Many teachers claim that a major reason for this situation is students' lack of interest in or passive resistance to tasks that challenge them to think and their desire to invest only a minimal amount of effort in academic work (Hampel, 1986; Tye, 1985). This contention seems to be supported by research indicating that in many high school classrooms students' covertly negogiate an agreement to exchange their compliant labor for minimal academic demands from their teacher (Powell, Farrar, & Cohen, 1985; McNeil, 1983, 1984; Sedlak, Wheeler, Pullin, & Cusick, 1986). Others (Goodlad, 1984; Sizer, 1984), however, suggest that students are indifferent to schoolwork because they are not intellectually challenged, but instead are confronted with a curriculum which emphasizes lower order cognitive tasks (i.e., the memorization of mindless information and the mechanistic application of algorithms).

One way to address this issue is to examine the perceptions of students regarding the kinds of academic work that make them think hard. This paper presents the findings of an empirical study which inquired into the characteristics of courses, lessons, and tasks that students report to be cognitively challenging, and their reactions to such challenges.

METHODOLOGY:

The data for findings in Chapters V and VI were collected as explained in the Introduction. A total of 45 students, nine from each of the five geographically diverse high schools, were interviewed for a total of approximately three hours on three different occasions spread over the 1986-87 school year. Interviews were tape recorded and students also were administered two questionnaires. Three students were selected from each of three different social studies classes in each school by the teacher on the basis of being representative of the top, middle and bottom third in the academic achievement distribution of the class. The composition of the sample of students was 47% male, 53% female, 24% high school freshmen, 29% sophomores, 31% juniors, and 16% seniors. By race, 73% of the sample were White, 22% were Black, and 4% were Asian.

Because of logistical constraints, the sample size is relatively small. A larger sample would have reduced the opportunity for in-depth interviewing needed in an exploratory study. However, since the students represented a cross-section with respect to academic

achievement (as borne out by the distribution of grade point averages), and the schools were located in diverse community contexts, the possibility of generalization should not be ruled out. Since both teachers and students were aware that the general purpose of the research was to examine HOT in social studies, it is possible student responses were biased toward favorable reactions to this subject. However, many students neither praised social studies nor their social studies teacher, and nearly all were able to give consistent reasons why certain classes and activities beyond the observed social studies class were challenging. In fact the consistency of student responses across a number of different learning contexts, at three different points during the school year, suggests the lack of a systmatic bias toward the observed social studies class.

In order to identify what students consider to be "challenging" classes and tasks, they were questioned about learning in several different contexts. Specifically they were asked to describe: (a) the most challenging course or subject they had taken both in the current school year and throughout high school and the reasons why it was challenging, (b) examples of challenging activities or tasks in their present social studies course that was observed by a HOT project researcher and their reactions (i.e., frustrations and satisfactions) to such activities, (c) specific questions or tasks from three social studies lessons observed by a researcher which made them think hard and the frustrations and satisfactions they experienced from thinking about those questions or tasks, (d) the types of lesson formats that generally challenge them to think hard, and (e) types of writing tasks that generally challenge them to think hard.

Three coding schemes were developed for categorizing students' verbal responses to the open-ended interview questions (and to follow-up probes by the researchers). Pilot schemes were constructed by generating categories based on relevant typologies suggested in the literature and the author's personal knowledge. Each scheme was reviewed by two other researchers from the HOT project and then tested on 10 randomly selected responses. After necessary revisions, the five schemes were subjected to a reliability analysis using a different set of 10 responses and two volunteer graduate students who were trained in coding the data. After reliabilities had been computed the raters (including the author) arrived at a consensus on the coding of disputed responses and the need for further revisions to some of the schemes. Using the revised schemes the author then coded the remaining 35 responses to each question.

FINDINGS:

The Most Cognitively Challenging Subjects

Over one third of the sample (see Tables 5A and 5B) rated a social studies course the most challenging class they had taken during the 1986-87 school year, and the most challenging they had taken throughout their high school career. The majority (i.e., 71%) of the social

studies courses mentioned as the most challenging class last year were history (compared with 60% of history courses in the project's sample), with the other five classes comprising two general social studies, two economics and one psychology. The proportion of social studies courses cited represents almost as many as the combined total for science and mathematics which are traditionally regarded as the tough subjects. Given the number of subject areas that these students have taken in high school (i.e., between six and ten), the expected probability (as a percentage) of social studies (or any one subject area) being rated the most challenging is between 10 and 16.7 percent.

It should be pointed out, however, that the five schools from which the subjects for this study were obtained were selected on the basis of having social studies departments which emphasized the development of higher order thinking. And in four of these five schools the principal rated social studies the strongest academic department in their school. Consequently, this finding of social studies being rated the most challenging subject cannot be considered representative of the general population of secondary school students. Nevertheless, it shows that in these schools many students find social studies particularly challenging.

Table 5A: Subject of Most Challenging Class Taken in Present School Year
(n=43)

Rank	Subject	Frequency	Percentage
1	Social Studies	17	35.4
2	Science	10	20.8
3	Mathematics	9	18.8
4	English	8	16.7
5	Foreign Language	2	4.2
6	Industrial Arts	1	2.1
6	Computer Science	1	2.1
		48*	100.0

* Five students nominated two classes as equally challenging

Table 5B: Subject of Most Challenging Class Taken in High School
(n=42)

Rank	Subject	Frequency	Percentage
1	Social Studies	15	35.7
2	Science	9	21.4
3	Mathematics	8	19.0
4	English	7	16.7
5	Industrial Arts	2	4.8
6	Foreign Language	1	2.4
		42	100.0

When asked : "What was it about this class that made you have to think hard?", half the students gave explanations that were related to cognitive processing demands (see Table 5C). The most frequent response can be classified as the demand for working out the answer or solution to what Simon (1973) has termed a "well-structured" problem where there is clear definition and goals and a right answer or solution. For example, one student compared working on trigonometry problems to a puzzle or game in which she gets satisfaction from working it out. On the other hand, other students reported that working on an "ill-structured" problem, where they had to formulate a response to or opinion on a question or issue in which the answer is problematic (and where the definition of the problem and the rules of solution are usually ambiguous), accounted for the reason a course was the most challenging. In the words of one student: "There were a lot of different ways to view the questions. Not direct ('yes' or 'no') questions, but rather 'what if. Usually there's more than one answer. You really had to think it through thoroughly to get to the best answer."

Interestingly, these two different types of problem were associated with distinctly different school subjects. Well-defined problems were the reason offered to explain the cognitive challenge of classes in mathematics, science, and in one case, auto mechanics. Ill-defined problems were mentioned only in reference to social studies. These differences were captured by one student who described chemistry and social studies as equally challenging but for different reasons:

> "Social studies is different from chemistry. Chemistry involves theories, formulas, and a lot of logic. There's only one or a few final answers. But it's a different kind of thinking than social studies. Social studies involves emotions and forming your opinion is different than thinking why something works. In social studies there is no one set answer, but a lot of factors involved in defending different solutions."

Table 5C
Reasons Class is Challenging
(n=43)

	no.*	%*	
comprehension	7	16.3	inherent difficulty of concepts
	5	11.6	student's lack of background in or knowledge of subject
27.9%	2	4.7	difficulty of reading material
work & memory load	8	18.6	quantity of work demanded
	4	9.3	demand for memorizing present material
	3	7.0	demand for accessing previous course material
25.6%	1	2.3	demand for independent work
	1	2.3	demand for concentration owing to uninteresting subject or material
cognitive processing demands	10	23.3	working out the answer or solution to well-defined problem
	5	11.6	deciding own position on question or issue with problematic answer
51.2%	5	11.6	making inferences
	3	7.0	applying multiple concepts or integrating multiple factors
	2	4.7	developing explanations, reasons and justifications for answers or opinions
	1	2.3	creating an original idea
teacher behaviors	2	4.7	lack of clear guidance in understanding material or defining task requirements
	2	4.7	demanding expectations and grading criteria
9.3%			
peer behaviors	3	7.0	competitive environment among students
	1	2.3	stimulating contributions from students
9.3%			

* Numbers add to more than 43 and percentages add to more than 100 as more than one category of reasons could apply to any response. Percentages were computed by dividing the number of students responding in each category by the total number of students (i.e., 43).

Another student perceived the same distinction between math and social studies, but found that the relativistic conception of knowledge conveyed in her social studies class made it much easier:

> "Social studies becomes routine after a while. You're expected to take a side and share your opinions. You can usually think of reasons to give and justify yourself. ....Calculus has right and wrong answers, but in social studies there is no right or wrong, usually. In social studies if you raise your hand and disagree with somebody it's okay."

When asked how she knows there aren't right answers in social studies, this student replied that her teacher told the class there are no right or wrong answers on opinion questions.

As Table 5D illustrates, solving ill-structured and well-structured problems were the most common reasons students gave for why social studies and other subjects respectively made them think hard. Almost one third of the students who reported social studies to be their most challenging class during the past year attributed the reason to their being required to decide their position or opinion on a question, problem or issue which has problematic answers. Two of these students specifically added that their teachers also demanded that they defend and support their opinions. One freshman student, for example, found her history class hard because "you have to defend yourself on every answer you give." On the other hand, one third of the students who found other subjects made them think the hardest were challenged by the task of figuring out the right answer or solution to a well-defined problem.

Table 5D: Most Common Reasons for Nominating Social Studies and Other Subjects as Most Challenging Class in Past Year

Rank	Social Studies	Fq*	%*	All Other Subjects	Fq*	%*
1	deciding own opinion	5	29	solving well-defined probs	10	32
2	making inferences	4	24	conceptually difficult	7	23
3	quantity of work	3	18	quantity of work	5	16

* Number and percentage of students

Some students reported that having to make inferences made a class challenging. All but one of these responses referred to a social studies class. These four students described how they were required in their social studies class "to go beyond the facts" or "to make inferences", by discussing, for example, the implications of particular decisions or events, or possible alternative scenarios and their

consequences. For example, one student stated that there are more underlying things he (the teacher) wants us to get. In earlier social studies classes "you'd read it over, in 1776 they signed the Declaration of Independence and we became a free nation, but now you have to think about it ... by going beyond the facts."

The same number of responses described the challenge of either applying multiple concepts/factors to a question or problem (e.g., "to do proofs, ..there are so many different possibilities to consider"), or developing explanations, reasons and justifications for answers or opinions (e.g., "She [the teacher] asks you why? She'll challenge you further"). In each of these two categories each reponse applied to a different subject. The example from the student who distinguished between the different kinds of thinking in chemistry and social studies also illustrates the challenge in social studies of integrating multiple factors given the variety of ways of viewing many social and historical questions.

In addition to demands for higher order cognitive processing, a significant minority reported that difficulties with comprehension made them think hard. This was mainly due to the complexity or abstractness of the concepts in a particular subject or their lack of background in or knowledge of the subject matter. This raises a second interesting distinction between social studies and other subjects. When students were asked why a particular school subject was the most challenging, the second most common reason offered was that the concepts involved are difficult to understand (see Table 5D). This reason was most often cited in the case of mathematics and science, especially physics, but never in social studies. Students don't seem to find the concepts and generalizations in social studies very difficult in themselves, although two students mentioned a problem in history of understanding events that occurred in a different time period to what they have experienced. Social studies apparently only becomes difficult when the teacher makes them work or intellectually play with the material. It is also possible that difficult or abstract concepts in the humanities can be connected more readily to concrete human events and experiences than concepts in science and mathematics.

These student perceptions seem to be consistent with the lack of a research tradition in the social studies field on how to teach and how students learn conceptually difficult material. Instead, the primary focus of curriculum theory and research in social studies historically has been the selection of content (i.e., identifying what to teach), which seems to be more problematic than in other subject fields.

In contrast to the demands for higher order thinking discussed above, many students indicated that the work and memory load in a class made it challenging, including the quantity of work covered, the amount

to be memorized, or the previous course material which had to be recalled. The quantity of work demanded was the third most common reason mentioned for classes in both social studies and other subjects (see Table 5D). Three students found the quantity of work covered in their history course very demanding, and five students said the same for either English or, in one case, Spanish, suggesting that the amount of material covered in these subjects is a challenge for many students.

Only a few students (see Table 5C) indicated that the academic environment created by their peers or specific characteristics of their teacher explained why a class was most challenging. In contrast, many students reported that the reason for a particular class being engaging was attributable to various characteristics of their teacher.

The Challenge of Higher Order Thinking. Several findings indicate that students find classes that attempt to promote higher order thinking challenging. Across all five schools, among students who nominated a social studies course as the most challenging in (a) the current school year, and (b) their high school career, the percentage who referred to a class which was observed as part of the HOT project was 76 (13 out of 17) and 73 (11 out of 15) respectively (see Tables 5E and 5F).

Table 5E: Number of Students in Each School Indicating Social Studies as Most Challenging Class Taken in Present School Year

Class	School					
	1	2	3	4	5	Total
HOT Social Studies	5	2	1	4	1	13
Other Social Studies	2	0	2	0	0	4
Total Social Studies	7	2	3	4	1	17
Other Subjects	3	10	6	5	7	31
	10*	12**	9	9*	8	48

* One student indicated two classes were equally challenging.
** Three students indicated two classes were equally challenging.

Table 5F: Number of Students in Each School Indicating Social Studies as Most Challenging Class Taken in High School

Class	School					
	1	2	3	4	5	Total
HOT Social Studies	4	0	2	4	1	11
Other Social Studies	1	1	2	0	0	4
Total Social Studies	5	1	4	4	1	15
Other Subjects	4	8	4	4	7	27
	9	9	8	8	8	42

The classes included in the HOT project's sample were selected by the social studies department head on the grounds that they emphasized higher order thinking (as well as contained a composition of students according to criteria established by the project staff). Observed classes were rated on 13 five-point scales, each of which focused on a different dimension of thoughtfulness. Six of these scales were later selected as the minimal criteria for judging the degree of thoughtful discourse displayed in a lesson (see Appendix 6). Every lesson received a score on these six criteria and mean ratings were computed for all the lessons observed in each school. Table 5G lists the rank orders among the five schools on: (a) their mean score on the minimal thoughtfulness criteria, (b) the number of students (out of 9) who indicated they were required to think very hard in the observed social studies class, and (c) the total number of students' nominating one of the HOT social studies classes as their most challenging class, either during the present year or throughout high school.

Table 5G: Rank Order of Schools on Minimal Thoughtfulness Criteria, Student Perceptions of Need to Think Hard, and Student Rating of Social Studies Class as the Most Challenging

School	Minimal HOT	# Students Think Hard	Most Challeng. Class
	RANK	RANK	RANK
1	3	1	1
2	4	4	4
3	2	1	3
4	1	3	2
5	5	5	4

In all three rankings the top three and the bottom two schools are the same, with schools number 2 and 5 consistently low on all three. And interestingly, the largest difference between schools on the

minimal HOT criteria was between the third and the fourth ranked schools. The (Spearman's Rank) correlation was 0.6 between the school ratings on the minimal HOT criteria and the number indicating that thinking hard was necessary for completing the work, and 0.55 between the minimal HOT ratings and the number of students' nominating a HOT social studies course as the most challenging class encountered this year or in high school.

Finally, two-thirds of the students indicated (on a five-point scale) that their present social studies course (the one identified by the HOT project) challenged them to use their mind more than other classes they had taken in high school (with one-fifth indicating a lot more). And only five students rated their present social studies class as less challenging. (The mean rating was 3.72 with a standard deviation of 0.98.) These ratings strongly support the conclusion that many students find higher order thinking cognitively challenging.

Activities and Tasks Perceived as Challenging

Slightly more than three-quarters (77.8%) of the students interviewed indicated that they had to think hard to complete the work in their present social studies course successfully. Of these students, all but three were able to recall and discuss an example.

When students' descriptions of activities in their social studies course that were challenging were categorized according to the four dimensional coding scheme (see Table 5H), one characteristic within each dimension tended to dominate. Almost two-thirds mentioned a writing task, almost three-quarters described tasks or activities performed alone, and over half spoke of an activity or task whose function was to manipulate or make sense of information. With regard to the fourth dimension of content, nearly two-thirds discussed tasks or activities in which the content involved either ideas or values.

Because of difficulties in the coding of this data on the content dimension (reliability of .62), student responses to this question were re-categorized using an alternative scheme. This scheme involved the four types of content challenges in social studies described by Newmann (in press), plus the categories of literal comprehension and recall. According to Newmann (in press), major cognitive challenges in social studies involve: (a) empathy or the incorporation of the experience of others into students' own thinking by trying "to see and feel the world from another's point of view"; (b) abstraction or the use of concepts (such as the nature of colonialism and the dynamics of global interdependence) or theories (such as the causes of economic depression) to try and "make sense of social events"; (c) inference or

Table 5H
Frequencies and Percentages for Students' Descriptions of Activities in Social Studies Course that are Challenging

type	no.	percent*
writing task	21	63.6
reading (only) task	3	9.1
discussion	3	9.1
debate, role-play	1	3.0
questioning	1	3.0
uncodeable	4	12.1
	33**	99.9

interaction	no.	percent*
self	24	72.7
peers	5	15.2
teacher	2	6.1
uncodeable	2	6.1
	33**	100.1

content	no.	percent*
ideas	13	39.4
values	7	21.2
specific topic	4	12.1
facts	4	12.1
uncodeable	5	15.2
	33**	100.0

function	no.	percent*
manipulation	22	53.7
creation	6	14.6
intake	5	12.2
demonstration	4	9.8
uncodeable	4	9.8
	41***	100.1

* Of total number of responses on each dimension from 32 students who described a challenging activity in their social studies course.
** One student described two challenging activities
*** More than one category of function applied to the description of seven activities.

going beyond the given data to develop explanations, to draw conclusions, or to make predictions; and (d) evaluation and advocacy or developing and defending "value judgements about what is good, right, and just in public life."

As Table 5I illustrates, the codeable responses were almost evenly spread among inference, evaluation and literal comprehension, with no examples of empathy and only a few of abstraction and recall. The high number of uncodeable responses was due to the lack of discussion by many students of the content involved in a challenging activity. This time the reliability, with two coders, was a more satisfactory 0.83.

Inferential challenges described by students included writing papers synthesizing the causes of the American revolution and explaining how Thomas Jefferson was stereotyped as a defender of democracy. Examples of evaluation were deciding whether or not the death penalty is justified and expressing and defending one's opinion on whether the motives of the framers of the Constitution were patriotic or selfish. Challenges in comprehension cited by two students were trying to understand psychological terms used in class and readings on Athenians which were written in their ancient language. Examining what makes a democratic nation and "coming up with a metaphor or simile for a culture" were two of the three examples of abstraction.

Turning to the types of tasks described as challenging, over a quarter of the 32 students who indicated that they had to think hard to complete their work successfully (and could recall an example) described an individual writing task which required manipulating (i.e., organizing, synthesizing, analyzing, etc.) ideas (i.e., the meaning of the material or principal ideas) or making value judgements. For example, a grade 11 student described writing a paper on causes of the American revolution: "We had to read several articles and bring the material together: economic causes, political causes. We had to state an underlying cause and how the other causes related to it."

In nearly all cases the writing tasks were completed at home or outside class, which may seem to suggest that homework assignments generally are regarded as more challenging than activities carried out in class. Only two students nominated activities involving direct interaction with the teacher. Yet four students described discussions or debates with peers as examples of being challenged (the fifth example of peer interaction involved a small group writing task).

Table 5I
Descriptions of Content of Activities in Observed Social Studies Course that were Reported to be Challenging
(n=32)

Type of Challenge	no.	percent*
inference	7	21.9
evaluation	7	21.9
literal comprehension	5	15.6
abstraction	3	9.4
recall	1	3.1
uncodeable	9	28.1
	32	100.0

* Percentages were computed by dividing the number of students responding in each category by the total number of students who reported a challenging activity (i.e., 32).

Note: 1. Ten students indicated that there were no challenging activities in their social studies course.
2. Three students said there were challenging activities, but they couldn't remember any examples.

These findings, however, may be biased by the wording of the question which stated: "Do you ever have to think really hard to complete your work successfully in this social studies course?". This question could have directed students' attention mainly to tasks which are graded, and in high school social studies that invariably suggests writing tasks done individually outside class. Interestingly, however, only four students discussed tasks whose primary function clearly was to assess their understanding of knowledge or skills acquired in the class. Tasks which demanded the creation of an original product, such as an original essay or visual metaphor, were more commonly mentioned, and most students' descriptions did not include any account of intentions or efforts to demonstrate knowledge to the teacher. But perhaps, given the wording of the question, this purpose was taken for granted and the students focused on other aspects of the task.

In contrast, when students were asked about the types of lessons that challenge them to think hard (see Table 4K which is discussed later), teacher directed activities were most commonly identified. This tends to support the above suggestion that a finer distinction may need

to be made in questioning students about the cognitive challenge of in-class activity versus homework. Unfortunately, no data was collected which would enable a comparison of the perceived amount of cognitive challenge of classroom activities directed by the teacher and academic tasks completed outside class. Approximately half the students in this study reported that a homework task in their present social studies course had made them think hard and more than half indicated that certain types of lessons controlled by the teacher also challenged them in general to think hard.

Reactions to Challenging Course Activities

All students who provided examples of when they had been required to think really hard to complete their work successfully in the observed social studies class, were asked to describe any frustrations and satisfactions that they experienced. Frustrations were categorized according to the coding scheme for reasons that an activity was challenging. While the specific types of frustrations that were recounted varied considerably, well over half the 29 students who experienced frustrations gave explanations that were related to demands for higher order cognitive processing (see Table 5J). No other major category of explanations for frustrations was applicable to more than five students.

Table 5J
Students' Reactions to Challenging Activities:
Frustrations
(n=29)

	no.*	%*	frustration from
comprehension 17.2%	2	6.9	difficulty of reading material
	1	3.4	student's lack of background in or knowledge of subject
	1	3.4	lack of reading material
	1	3.4	comprehension but specific reason unclear
work & memory load 13.8%	3	10.3	quantity of work demanded
	1	3.4	demand for concentration owing to uninteresting subject or material
	1	3.4	insufficient studying (for test)
cognitive processing demands 58.6%	4	13.8	finding correct/relevant information
	3	10.3	deciding own position on question or issue with problematic answer
	3	10.3	developing explanations, reasons and justifications for answers or opinions
	2	6.9	creating an original idea
	2	6.9	making inferences
	2	6.9	organizing information
	1	3.4	cognitive processing but specific demand unclear
teacher behaviors 10.3%	1	3.4	demanding expectations and grading criteria
	1	3.4	fear of failing or not performing well
	1	3.4	teacher doesn't understand student's thinking
peer behaviors 3.4%	1	3.4	peers successfully attacking your arguments
uncodeable	1	3.4	

* Numbers add to more than 29 and percentages add to more than 100 as more than one category of reactions could apply to any response. Percentages were computed by dividing the number of students responding in each category by the total number of students who reported a frustration (i.e., 29).

There was much more commonality in the case of satisfactions which were categorized using the coding scheme for reasons that an academic activity was engaging (see Table 5K). Two-thirds of the 30 students who reported receiving some form of satisfaction from thinking hard were intrinsically rewarded by either their performance in completing the task or the value they attached to the task itself. Of the former group, many described their rewards as the feeling of having done a good job or of having performed well on the particular task, while the rewards for others came from successfully completing the task.

These two forms of satisfaction may appear very similar, but the first refers to a strong sense of satisfaction with the quality of one's work and the effort one has made. For example, one student described how, after spending 10 hours writing a very diffficult book review, she read her completed work and felt "happy with it .. I got satisfaction from a job well done." Whereas the second involves a feeling of being pleased with having accomplished the set task regardless of the quality or extent of one's effort, as exemplified by the student who said: "I succeeded in debating something I didn't believe." It is possible (as demonstrated by the responses of two students) to experience both these satisfactions from the one task. Learning or thinking through something successfully on one's own was mentioned as the intrinsic reward by two students.

Several students mentioned the satisfaction of developing new insights or better understandings. As one student emphasized:

> "I get more satisfaction from truly understanding the material and having it organized in my mind than from doing a really good report. I get a lot of satisfaction from really, really knowing the material and essays are wonderful for helping you ....to really understand the material" (due in part, she explained, to researching different sources).

Others felt good about either having to think hard or just being able to think for themselves. One of these students described the satisfaction he received in English from not having to worry about what the teacher thinks when he disagreed with his teacher's interpretation of the symbolism in a story.

Table 5K
Students' Reactions to Challenging Activities: Satisfactions
(n=30)

	no*	%*	satisfaction from
instrumental performance on task 26.7%	5	16.7	good grade or comment from teacher
	2	6.7	finishing task (sense of relief)
	1	3.3	performing better than others
	1	3.3	passing the test or assignment
instrumental value of task 6.7%	1	3.3	acquiring knowledge or skills that can be used outside school
	1	3.3	acquiring information that can be used on a future assignment
intrinsic performance on task 46.7%	8	26.7	performing well or doing a good job
	6	20.0	successfully accomplishing task
	2	6.7	learning or thinking successfully on my own
	2	6.7	intrinsic performance but specific reason unclear
intrinsic value of task 40.0%	7	23.3	developing new insights or better understandings
	3	10.0	having to think hard
	2	6.7	being able to think for myself
	1	3.3	creating original or unique product
	1	3.3	sharing understandings with others
uncodeable	2	6.7	

* Numbers add to more than 30 and percentages add to more than 100 as more than one category of reactions could apply to any response. Percentages were computed by dividing the number of students responding in each category by the total number of students who reported a feeling of satisfaction (i.e., 30).

Note: 1. Twenty students (66.7%) mentioned intrinsic (i.e., performance or value) satisfactions, and ten students (33.3%) mentioned instrumental (i.e., performance or value) satisfactions.

2. Only six students gave reactions of satisfaction that were applicable to more than one broad category (e.g., intrinsic performance and instrumental performance).

Besides these intrinsic rewards, eight students admitted receiving satisfaction from extrinsic rewards that resulted from their performance on a task. For five of these students the source of reward was a good grade or words of praise from the teacher. And somewhat surprisingly, no student - even in the competitive upper track classes - spoke of any sort of recognition from peers as a source of satisfaction, and only one mentioned being rewarded by doing better than his classmates. Satisfactions derived from the instrumental value of a task were not common.

Although the students were invited to talk about as many frustrations as they could remember, the vast majority mentioned only one. In contrast, most students described more than one type of satisfaction that they derived from thinking really hard. Among those who offered more than one explanation or description of frustration, there were only two students whose responses applied to two different major categories. This time a similar trend was evident for satisfactions where the responses of only six students fell into more than one of the four major categories. In other words, for 80 percent all their satisfactions were related to <u>one</u> of the following: instrumental performance, intrinsic performance, instrumental value or intrinsic value of the task. But more importantly, only 20 percent of students derived satisfaction only for instrumental reasons.

<u>Challenging Questions or Tasks in Observed Social Studies Classes</u>

In addition to general questions about school subjects and the social studies course included in the HOT study, students were also asked on each of the three visits to reflect on a particular social studies lesson observed by a researcher. The responses to the question: "What questions or tasks in that lesson, if any, made you really think or use your mind?" did not fit the categories used for analyzing students' more general descriptions of challenging activities in social studies. So, instead, this data was categorized solely by relying on the alternative scheme employed for coding the content component of challenging social studies activities, namely the extension of Newmann's (in press) four types of challenges.

As Table 5L indicates, over 80 percent of the responses could be categorized within one of the four content challenges that demand higher order thinking (i.e., inference, evaluation, abstraction, and empathy). The most frequently mentioned examples were of inferential challenges where students had to go beyond the data to draw conclusions. For example, one student described trying to figure out what England or France really did, if anything, to develop their colonies, while another student was challenged by the self-initiated task of explaining why the standards and quality of food we enjoy today

are so different from those during the time of the meat packers' controversy.

Table 5L
Descriptions of Questions or Tasks in Observed Social Studies Lessons that were Reported to be Challenging
(n=99)

Type of Challenge	no.	percent*
inference	34	34.3
evaluation	26	26.3
abstraction	21	21.2
recall	6	6.1
empathy	4	4.0
literal comprehension	2	2.0
uncodeable	6	6.1
	99	100.0

* Percentages were computed by dividing the number of student responses in each category by the total number of responses (i.e., 99).

Note: 1. In 22 cases students indicated that there was no challenging question or task in an observed lesson.
2. Twelve students were not taking the same social studies course (i.e., the one included in the HOT sample) on the third visit and hence were questioned on only two observed lessons.
3. Three students were absent from the observed lessons on one visit and hence were questioned on only two lessons.
4. More than one type of challenge applied to one response.

Consistent with the findings already discussed on the reasons students find social studies classes to be challenging, a number of responses referred to questions or tasks which required making a valuejudgement. Examples cited by students included deciding whether civil disobedience is justified, whether Jefferson acted properly in establishing an embargo, and whether nullification could ever be justified.

When identifying challenges in specific lessons, students described problems of abstraction (in making sense of a concept or theory) more frequently (one in five) than when discussing challenges of the social studies course (one in ten) or of school subjects generally (no mention of conceptual or abstraction problems). Distinguishing between communism and socialism, understanding the

checks and balances of presidential and congressional power, and defining "ethnic pluralism" are some of the challenges in abstraction reported by students in these lessons. A possible explanation of why no students mentioned abstract concepts as a reason for finding social studies a challenging subject is that these challenges may be irregular and short-term, rather than a continuing cognitive challenge as might be the case for developing and defending explanations or opinions.

Only a few students spoke of having to think hard to gain an empathetic feel for experiences occurring in a different time and/or place. Specific examples offered were: trying to understand why security guards in a prison have certain attitudes toward inmates, "understanding the concept of an historian's views being influenced by the time in which they wrote", and trying to see "the way the elite see and the way the farmer sees things" (at the Constitutional Convention). The low number of empathy challenges could be attributable to teachers not presenting tasks aimed at this objective in a challenging form, but instead using films, stories or lectures for this purpose which may not require students to actively develop answers to challenging questions.

In addition to the above higher order thinking challenges, only a few responses indicated that lower order cognitive tasks were challenging. The majority of these involved the recall of previously encountered information, with the remainder being concerned with comprehending the literal meaning of new information.

Only two students found nothing challenging in any of the three observed lessons on which they were questioned, three students reported nothing challenging in two of the three lessons, and another 10 said there was no question or task that made them think hard in one of the observed lessons. And interestingly, no student described the same type of challenge in all three lessons.

Reactions to Challenging Questions or Tasks in Observed Lessons

As in the case of challenging course activities, students who indicated that a question or task from a lesson observed by a researcher made them think hard, were asked to describe any frustrations and satisfactions they encountered in responding to that particular question or task. It is interesting to compare the responses to this question with those to the equivalent question on reactions to challenging course activities. Such a comparison provides a check on the consistency of student responses and an indication of the validity.

The percentage of responses in each of the major categories of frustrations from dealing with challenging questions or tasks (see Table 5M) was very similar to those for challenging course activities.

Table 5M
Reactions to Challenging Questions or Tasks
in Observed Social Studies Lessons: Frustrations
(n=47)

	no.*	%*	frustration from
comprehension 17.0%	3	6.4	lack of sufficient information
	2	4.3	difficulty of reading material
	2	4.3	student's lack of background in or knowledge of subject
	1	2.1	comprehending literal meaning of concept
work & memory load 6.4%	2	4.3	demand for accessing previous course material
	1	2.1	demand for memorizing present material
cognitive processing demands 48.9%	8	17.0	deciding own position on question or issue with problematic answers
	5	10.6	having to make abstractions (to make sense of information)
	3	6.4	developing explanations, reasons and justifications for answers or opinions
	3	6.4	making inferences
	2	4.3	finding correct/relevant information
	2	4.3	taking the perspective of another person
teacher behaviors 6.4%	1	2.1	lack of clear guidance in understanding material
	1	2.1	uncertainty about teacher's response
	1	2.1	lack of opportunity to create own solutions
	1	2.1	insufficient time for discussion
peer behaviors 10.6%	3	6.4	trying to convince peers of your argument
	1	2.1	peers quicker at answering questions
	1	2.1	disagreements in peer group
uncodeable	5	10.6	

* Numbers add to more than 47 and percentages to more than 100 as more than one category of reactions could apply to any response. Percentages were computed by dividing the number of student responses in each category by the total number of responses (i.e., 47).
of the interview schedules.

Ten percent fewer students reported frustrations related to higher order cognitive processing demands in observed lessons, although this category still represented approximately 50 percent of the students. Within this category, more students indicated frustrations in observed lessons (than in the whole course) with deciding their position on a problematic issue. This may be the result of dealing with new and conflicting information in a lesson, whereas assignments to a large extent involve writing about an issue that has already been discussed and that students have had some time to think about. As one student explained in discussing a class argument on the issue of civil disobedience, "it was a little frustrating - cause there's no real answer. Part of it may be because since it's the first time we've discussed it, it's a little confusing."

On the other hand, fewer students encountered frustrations with finding correct or relevant information in the observed lessons. This is not surprising as students' examples of having to think hard to complete work successfully in the social studies course generally involved writing papers, which, of course, require searching for appropriate information.

The number of students who described satisfactions derived from their intrinsic performance was noticeably less - approximately one-third instead of half - on questions or tasks in the observed lessons than on examples of assignments done throughout the course (see Table 5N). But we would expect more satisfaction when the students can choose an example of the most challenging activity or assignment encountered over the span of the entire course. Their performance on such a major and demanding activity is more likely to create intrinsic satisfaction than on just a question or task from a lesson selected at random. Nevertheless, the two most common descriptions of satisfaction were the same in both cases, namely: successfully accomplishing the task, and performing well or doing a good job.

This last description, however, did drop from being mentioned by 27 percent of the students in relation to course activities to only eight percent in relation to observed lessons. A student who felt he did a good job on a question in an observed lesson explained that

> "it was satisfying to know what I finally believe. When I decide based on considering evidence and different points of view, you know you've made a good decision because you've seen both views .... The reward is when, after you reach a conclusion, you can articulate your position, knowing you've examined all sides of the issue."

Table 5N
Reactions to Challenging Questions or Tasks in Observed Social Studies Lessons: Satisfactions
(n=63)

	no*	%*	satisfaction from
instrumental performance on task 27.0%	7	11.1	good grade or comment from teacher
	6	9.5	performing better than others
	2	3.2	receiving recognition from peers
	1	1.6	teacher failing to prove me wrong
	1	1.6	seeing that peers think the same as me
instrumental value of task 1.6%	1	1.6	acquiring knowledge or skills that will be useful in the future
intrinsic performance on task 31.8%	13	20.6	successfully accomplishing task
	5	7.9	performing well or doing a good job
	3	4.8	learning or thinking successfully on my own
	2	3.2	being able to convince others of my argument or defend my opinion
intrinsic value of task 34.9%	15	23.8	developing new insights or better understandings
	4	6.3	having to think hard
	2	3.2	sharing understandings with others
	2	3.2	hearing appealing ideas
	1	1.6	being able to build further ideas
uncodeable 11.1%	7	11.1	

* Numbers add to more than 63 and percentages to more than 100 as more than one category could apply to any response. Percentages were computed by dividing the number of student responses in each category by the total number of responses (i.e., 63).

Note: 1. Fourty two student responses (66.7%) mentioned intrinsic (i.e., performance or value) satisfactions, and 18 student responses (28.6%) mentioned instrumental (i.e., performance or value) satisfactions.

2. Only four students gave reactions of satisfaction that were applicable to more than one broad category (i.e., intrinsic performance and instrumental performance).

Such an examination, however, is generally more likely to occur through the process of completing a major assignment than dealing witha question or task in one relatively brief lesson. Thus, the predominance of assignments as the examples of challenging course activities is sufficient to explain the greater occurrence of satisfaction from performing well in this context compared to that derived from challenging questions or tasks in a lesson.

On the other hand, developing new insights or better understandings, such as "it gave me more knowledge of the different categories in which the psychologists' approaches could be placed", remained a source of satisfaction for almost one quarter of the student reponses. Even if students only "understand it (how early man survived) a little better" after an in-class task, they at least feel they "can say you learned something today." And while most students' insights related to a new or increased understanding of content, some experienced more far reaching insights about people and their views:

> "It showed me that people have different views and that people think differently... So now I try to have a more open view because different people interpret in different ways. So maybe it will make me a more open person to what people think."

In addition to intrinsic value of the task, there was also a noticeable similarity in the percentages of students on the other two major categories of satisfaction (i.e., instrumental performance and instrumental value) between the reactions to challenging course activities and the reactions to challenging lesson tasks. This consistency confirms that approximately one third to one half of the students in this study derive satisfaction from their intrinsic performance, more than one third from the intrinsic value of the task, and only approximately a quarter from their instrumental performance on a challenging task.

Types of Lesson Formats that Challenge Students to Think Hard

Students were requested to indicate the general types of lesson formats (in all subjects) that challenged them to think hard (see Table 50). Two teacher directed activities, teacher-led class discussion and questioning class on understanding of subject matter, had the highest frequency of checks. In both cases, students explained the need to concentrate on understanding the discourse so that they would be prepared to answer any questions if called upon. As one student explained, "you have to be thinking all the time when you have a teacher who doesn't just call on volunteers." Furthermore, two other students emphasized that "teachers who make you think the most keep picking on you if you don't give a reasonable answer."

Table 50: Types of Lessons Students Report as Challenging Them to Think Hard
(n=43)

Rank	Lesson Format	Frequency	Percentage
1	class discussion led by teacher	21	48.8
1	teacher questioning class on understanding of subject matter	21	48.8
3	working on a task by myself	19	44.2
4	class discussion led by students	12	27.9
4	depends on the class or teacher	12	27.9
6	lecture	9	20.9
7	working on a task in a small group	8	18.6
8	working on a task with a friend	6	14.0
9	film	3	7.0
10	other	2	4.7
11	none	1	2.3

Apparently, teacher questioning creates pressure by putting students "on the spot", "whereas in class discussions led by students you don't have to say anything" or "you can just say what you think." Consequently, although ranked fourth, only just over one quarter of respondents find class discussions led by students challenging in contrast to the two-thirds who indicated that they are interesting and worthwhile. The challenge for these students is usually in role plays or debates where they have to develop arguments "for something you don't really believe in" or "to convince others about your opinion."

Closely behind teacher-led questioning and discussion was working on a task by yourself when, according to a number of students, all the thinking had to be done on your own. Not only do you have to get along "without the benefit of others' ideas and views," but also "there's no feedback when you work alone." On the other hand, working on a task in a small group or with a friend is far more engaging (rated so by more than half) than challenging (as indicated by less than one in five). Several students admitted that groupwork was fun, but often involved discussing non-task matters. But the major reason advanced for group tasks being less challenging was the opportunity they provided to rely on other people to do the thinking. Yet some students found it challenging in a small group task to connect up different people's ideas, while others suggested that overcoming the inevitable disagreements and arguments (over procedures as well as substance) was always a challenge.

Lectures were rated engaging and challenging by roughly equal percentages, but only three students thought a film made you think hard compared to 18 students who rated it interesting and worthwhile. Lectures are hard for some students because "some things go by you, but the teacher can't stop each time something is covered that you don't understand." Said another way by another student, there is not the opportunity for things to be repeated.

Many students indicated that the type of lesson which makes you think hard depends on the subject or teacher. Again, one interesting distinction was between subjects where there are right answers and those involving opinions or value judgements:

> "If one thing is right - as in math and science - you have to think harder. That is, trying to rationalize F=MA is harder than figuring out my opinion on the progressives."

Types of Writing Tasks Found Challenging

As well as the different types of lesson or instructional formats, students' reactions also were sought to the challenge of different types and different objectives of writing tasks. Almost 80 percent indicated that writing a paper challenged them to think hard (see Table 5P). Percentages then declined as the type of writing decreased in length or substance. Thus, writing paragraphs or a short essay was challenging for about half the students, making an outline the same for one in five, and writing one or two sentences made less than one in 10 think hard.

Table 5P: Types of Writing Students Report as Challenging Them to Think Hard
(n=43)

Rank	Type of Writing	Frequency	Percentage
1	writing a paper	34	79.1
2	writing paragraphs or a short essay	22	51.2
3	depends on the topic or subject	14	32.6
4	making an outline	9	20.9
5	taking notes (from a lecture or book)	7	16.3
6	writing one or two sentence answers	4	9.3
7	other	3	7.0
8	none	2	4.7

Generally students report that the major challenge of writing papers is finding enough information, especially on topics that they

feel they know little about. And not only is sufficient information required, but "you have to think of different ideas that will bind together ... and that will support your thesis." Another difficulty mentioned by several students is organizing the paper, for example "into a topic sentence, logical order of paragraphs, introductory and concluding sections."

Even a couple of paragraphs is tough for some students, one of whom candidly stated: "You get tired of writing so you just put down anything. A paragraph isn't too hard, but anything longer is hard to find enough things to write down." Yet for other students shorter writing tasks are harder, because of the need to be succint and to "fit all your ideas into a short space." One student, who has "lots to say", finds it hard to use only specific facts, to omit unimportant details, and to not be too general.

Making an outline provides the challenge apparently of deciding what to include and the order in which to include it, while taking notes from a lecture (but not a book) demands thinking "because the teacher won't stop and you have to pick out important points quickly." And, of course, you have to be "able to read your own writing afterwards."

Many students indicated that any of the listed writing tasks can be challenging depending on the subject or particular assignment. For one student, "the organization of other people's ideas that you learn is easier than expressing your own thoughts." Somewhat in contrast was the student who found papers in social studies hard because the topic is always one that is just being learned, whereas in English often the information is already known (e.g., your experiences).

In response to the question concerning the nature or objective of writing assignments that challenge you to think hard, five of the six listed were checked by a third or more of the students (see Table 5Q). Writing which involved taking a position on an issue and defending it, explaining a concept/principle/ theory/problem or issue, and reporting on independent research were each rated as hard by approximately half the students. The first was generally said to be challenging because you have to think of reasons and "you may like something but not know how to defend it." Or even if you do have reasons, "you have to be sure of what you are defending or opposing and that causes you to prepare the utmost to cover all the bases so you will not be caught off guard." Explanations of concepts, principles, etc. were reported to require thinking in order to put them in your own words, especially when the "sources are in complex terms you have to simplify." The most descriptive explanation of the cognitive demands of report writing involved the need to include both sides, or the pros and cons, of the positions on the issue being researched: this meant "you have to continually think whether you have covered everything."

Table 5Q: Subject of Writing Assignments Students Report as Challenging Them to Think Hard
(n=43)

Rank	Type of Writing Assignment	Frequency	Percentage
1	taking a position on an issue and defending it	23	53.5
2	explaining a concept, principle, theory, problem or issue	21	48.9
3	reporting on independent research	20	46.5
4	creative writing (making up stories or poems)	18	41.9
5	describing my experiences or feelings	14	32.6
6	depends on the topic or subject	10	23.3
7	summarizing material covered in class or readings	8	18.6
8	other	2	4.7
9	none	1	2.3

Creative writing is hard for many students owing to either the demand to be original or the lack of concrete material with which to work: "you start from scratch and make it up by using your imagination," whereas in other forms of writing "you have something to start with." Experiential (i.e., describing experiences or feelings) writing, as described by one student, demands getting "down to your feelings, looking at how you really feel, it's not that easy." But the concern of another student was revealing too much of his feelings "because it gets into the wrong hands." Unfortunately, for this student making decisions about "what to tell" is the hard part.

By contrast, one student never found creative writing hard "as it comes from within you. You're also not concerned about whether you have got things right or wrong as you know you're not going to be graded on facts or what you have learned." The emphasis on giving (right) reasons in students' explanations of the challenge of expository writing suggests that others may not find creative writing challenging for this same latter reason.

Trying to identify the most important facts seemed to be the main challenge for the few students who indicated that summarizing material made them think hard. Typifying many of the reasons offered for why the challenge of writing assignments depends on the topic or subject was the particularly optimistic reponse that:

> "Anything that's worthwhile challenges you to think. Creative writing allows you to choose from unlimited options. Writing a

story involves thinking of an original plot, and that's hard. Taking a position challenges you to articulate your thoughts that are often fuzzy to begin with. You have to understand to explain a concept, and this forces you to think."

SUMMARY:

Over a third of the students interviewed in this study nominated social studies as the most intellectually challenging subject they had taken in high school, while a few more said that science or mathematics was the most challenging, and one in six nominated English. More than half the students who selected social studies gave the reason as either having to address ill-defined problems or questions with problematic answers, or being required to make inferences. The other half were equally divided in their reasons between having to overcome comprehension difficulties and coping with the work or memory load, especially the quantity of work demanded. By contrast, half the students challenged the most by science/math attributed the reason to solving well-defined problems with single, correct answers. Most of the others found the concepts in science or math difficult to understand. The quantity of work was the contributing reason provided by half the students reporting English to be the most challenging subject.

Two findings strongly suggest that many students find higher order thinking challenging. First, a large number of students indicated that the course included in our study of social studies departments emphasizing higher order thinking was their most challenging class. Second, there was a reasonably high correlation between school rankings on the number of such courses nominated as the most challenging and the rankings on the extent of higher order thinking observed in classrooms.

Three in four students indicated that they had to think hard to do their work successfully in social studies. The majority mentioned individual writing tasks completed outside class as examples of a challenging activity. The type of content challenge selected for social studies courses in general tended to involve inference, evaluation or literal comprehension. Inferential and evaluation challenges also were the most common types described by students in relation to questions and tasks that made them think hard in a lesson observed by a researcher, and problems of abstraction were also quite often noted in particular lessons.

Students were divided in their ratings of the degree of difficulty encountered when their social studies teacher posed a problem or question where there was no single correct answer. Approximately half rated such problems quite difficult, while the other half rated them not difficult. This is consistent with the findings on reasons for different subjects being challenging where some students selected social studies because of its ill-defined problems with problematic answers, while others selected another subject (usually science or mathematics) because of its well-defined problems which they found more challenging.

Both teacher-led class discussions and teachers questioning members of the class on their understanding of the material reportedly made almost half the students think hard because of the need to be prepared to answer a question when called upon. Almost as many students indicated that working in class on a task by oneself was challenging: for some because of the lack of input or feedback from others.

The vast majority of students reported that writing a paper was challenging, usually because of the amount of information that had to be found. In general, the shorter the type of writing assignment the fewer the number of students who found it made them think hard. As for different types of writing assignments, taking a position on an issue and defending it, explaining a conceptual-theoretical material, and reporting on independent research, were each checked as making them think hard by approximately one in two students.

The Conclusion to Chapter VI will relate findings on student perceptions of cognitively challenging work to their perceptions of engaging academic work.

VI

STUDENT PERSPECTIVES ON ENGAGEING CURRICULUM

Robert B. Stevenson

INTRODUCTION:

A common theme in much of the contemporary literature on secondary schools is that students generally are indifferent to schoolwork and exert a minimal amount of effort (Boyer, 1983; Cusick, 1983, 1973; Everhart, 1983; Goodlad, 1984; Powell, Farrar, & Cohen, 1985; Sedlak et al, 1986; Sizer, 1984). Teachers strongly agree, rating student passitivity and lack of interest as their worst problem (Hampel, 1986; Tye, 1985). This academic apathy or low-level participation in schoolwork, however, is a symptom of a deeper condition of student disengagement (Natriello, 1984).

As described in the recent literature, the problem of disengagement in rigorous academic work afflicts students of all levels of academic ability and from all socioeconomic and racial classes (Goodlad, 1984; Sizer, 1984; Cusick, 1983; McNeil, 1983). Although the higher drop-out rates for Hispanics, blacks, and youth from households of lower socioeconomic status (U.S. General Accounting Office, 1986) suggest that students from such backgrounds are more likely to have strong negative attitudes towards school and to be alienated from the institution, lack of academic engagement is certainly not confined to alienated students in general or these students in particular. In fact, according to a recent report there is a long tradition of the vast majority of high school students being indifferent to academic learning, with the problem claimed to have become exacerbated in the last twenty years (Sedlak et al, 1986). The lack of engagement among most students is manifested in a more neutral attitude where schoolwork arouses neither negative nor positive feelings and a passable but not utmost effort.

Although disengagement has been described extensively in the literature on secondary schools, where the problem seems to be most serious, the authors of a recent review concluded that there appears to be no study that has examined the other end of the continuum: student engagement in academic work (Mosher & MacGowan, 1985). Yet, in order to tackle the problem of disengagement, it is critical to understand what engages high school students in academic work. More precisely, we need to identify the conditions that will facilitate a positive attitude and the exertion of a committed or maximum, rather than apathetic or minimal, effort.

Given the lack of research on this subject an appropriate starting point would seem to be the perceptions of students regarding what kind distinguishing characteristics of some high school classrooms (as the settings for academic work) which students associate with engagement?

A number of studies have suggested that schools (i.e. classrooms) vary considerably in the extent of student engagement (Rutter et al, 1979; Wehlage et al, 1980; Wynne, 1980). Therefore, this study sought to identify the characteristics (e.g. curriculum content, format, instructional organization), as perceived by students, of lessons and academic tasks which generally evoke engagement in thinking and learning. For this purpose student engagement in academic learning is defined as a learning situation (i.e. a class lesson or academic task) in which: a) the student makes a serious or committed effort to master the knowledge and/or skills intended to be developed, and b) the student values the work itself or the actual process and/or substantive outcomes of learning (rather than the institutional rewards which might accrue) as meaningful and worthwhile, or finds the topic intrinsically interesting.

This chapter presents the findings in relation to the two research questions:

(a) What kinds of academic work do high school students report as being engaging? and

(b) What reasons do they give for finding such work engaging?

Students were questioned about "engaging" academic work in several different learning contexts. These contexts consisted of any subject or course the student nominated as particularly engaging, the social studies course that was observed by a HOT project researcher (both in general and within specific observed lessons), and any class where the student had experienced in-depth work. Specifically, students were asked to describe: (a) the most engaging course or subject they had taken in the current school year, (b) examples of the most engaging activities or tasks from a particularly engaging class (in any subject), (c) examples of engaging lessons or assignments in their present social studies course, (d) specific questions or tasks from three social studies lessons observed by a researcher which they found engaging, (e) anything, if at all, they found engaging about any in-depth work they had experienced in school, (f) teacher behaviors associated with engaging class(es), and (g) types of lesson formats they generally found engaging.

In all cases reasons were sought from students as to why the examples they provided were engaging. It was from these reasons that we hoped to learn the most about students' perspectives on the characteristics of engaging academic work.

Most Interesting and Worthwhile Course

When asked what was the most interesting and worthwhile class they had taken during the present school year, almost half of the students indicated a social studies course (see Table 6A). The social studies courses described as the most engaging comprised 10 history, three general social studies, two economics, two psychology, and one law. English, including two speech classes and a writing class, was the only other subject area nominated by more then ten percent.

The students interviewed, however, came from five schools that were selected for the attention their social studies departments gave to fostering higher order thinking (HOT). Furthermore, in four of these schools the principal considered that department to have the best teachers. So, while these schools are not representative of high schools in general and therefore the social studies experiences of their students is not representative, these findings do indicate that many students find social studies particularly engaging. This finding is particularly encouraging given that other studies have found that social studies is rated low in interest level by secondary students: for example, in one study only 13 percent indicated that social studies was their favorite subject (Shaughnessy & Haladyna, 1985).

Students were asked to explain the reasons why they found the course they nominated engaging. Although engagement has been defined in this study in intrinsic terms (i.e., the subject matter, or the process or outcomes of learning are valued for their own sake or regarded as intrinsically interesting), students were asked to indicate a course that was interesting and worthwhile. In other words, as students

Table 6A: Subject of Most Engaging Class Taken in Present School Year
(n=43)

Rank	Subject	Students Frequency	Percentage
1	Social Studies	18	41.9
2	English (incl. Speech)	8	18.6
3	Science	4	9.3
4	Mathematics	3	7.0
4	Foreign Language	3	7.0
6	Industrial Arts	2	4.7
6	High Technology Lab	2	4.7
8	Art	1	2.3
8	Health	1	2.3
8	Photography	1	2.3
		43	100.1

explained the reasons for their engagement they could characterize classes as "interesting" for either extrinsic or intrinsic reasons, or as "worthwhile" for instrumental or non-instrumental reasons.

As Table 6B illustrates, over three-quarters of the students gave at least one reason that referred to the intrinsic value of the class (i.e., they were engaged by either the content, processes or outcomes of learning in that class). The most common single reason in this category (mentioned by nearly half the students) was that the subject matter was intrinsically interesting, often owing to its perceived relevance to the real world or, more specifically in the case of social studies, to current issues and events. Chemistry, for example, was described as engaging because "it answers questions you come across in day to day life: for example, why is frozen ice less dense than water?" One student found U.S. History engaging because "the material relates to our current political and economic system." Yet for another student, history was intrinsically interesting, not for its contemporary relevance, but because "we deal with very informative topics and real characters whom you can relate to - you can relate to what happened, how people felt and what they went through, for example, in colonial times." The particular topics that were mentioned as intrinsically interesting in any subject tended to be idiosyncratic with no topic engaging several students.

The other common reasons in this category of intrinsic value involved the opportunity to actively participate in the class, by discussing one's own opinions or ideas, by creating an original product, by manipulating information or materials (e.g., technical drawing and scientific instruments), or by working on a challenging task. These four categories combined represented a reason given by more than two out of every five students questioned. The engagement power of being able to express opinions or ideas in class is conveyed by the following three quotations from students:

> "Every week we had maybe two speeches to write and everything....they were topics that everybody could relate to....like favorite teachers and it gave us the opportunity to say what we have to say about teachers, about what we didn't like about school ...things you normally don't have the opportunity to say."

Table 6B
Reasons Subject or Course was the Most Engaging
(n=43)

	no.*	%*	
instrumental performance in class 2.3%	1	2.3	subject was easy to learn
instrumental value of class 14.0%	6	14.0	relevant to future goals
intrinsic performance class 7.0%	3	7.0	performed well
intrinsic value of class 79.1%	19	44.2	intrinsically interesting subject matter
	10	23.3	opportunity for discussion or contribution of own ideas/opinions
	5	11.6	learned new things or gained new insights
	4	9.3	opportunity to manipulate information or materials
	4	9.3	opportunity to create unique product
	2	4.7	cognitively challenging task
	2	4.7	opportunity to hear ideas/opinions of others
	1	2.3	opportunity to share understandings with others
	1	2.3	logical subject matter, makes sense
teacher behaviors 41.9%	9	20.9	instructional practices (made subject interesting, challenging)
	5	11.6	attitude towards students (caring, gave responsibility)
	5	11.6	personal characteristics (likeable, sense of humor, enthusiastic)
	2	4.7	demanding academic expectations
	1	2.3	knowledgeable about subject
	1	2.3	provided variety (of activities)
	5	11.6	specific teacher behavior unclear
peer behaviors 2.3%	1	2.3	serious attitude to learning

* Numbers add to more than 43 and percentages add to more than 100 as more than one reason could apply to any response. Percentages computed by dividing number of students responding in each category by total number of students (i.e., 43).

"The teacher lets us do more, he gives us more freedom so that it feels more like high school should be, whereas other classes seem more like junior high. He starts us off and then the class teaches the class, but he corrects us when we are wrong or need help so that we are not left alone."

"Everyone can say what they want to say. In large groups you can say what's on your mind. You can say at the end of the class that you contributed to the answer."

Similar adolescent needs seem to be met when students are able to create a unique or original product, such as showing "mood in a plaster mask" or using "orange coral and sea plants with orange fish" in a printing in an art class. Another form of participation that apparently is engaging is mental or physical manipulation, such as solving problems or applying concepts (e.g., in mathematics "there's lots of figuring, concepts to understand and apply"), or experimenting with equipment (e.g., "it was fun to work with the voltmeters and experiment with different voltages and light bulbs and stuff like that.").

Besides intrinsic interest and opportunity to participate, five students reported that developing a better understanding of the material or gaining new insights was their reason for being so engaged by a course. An example of the former case is: "the class clarifies what you have already read at home. It clears up a lot of confusion about it [our political and economic system]." And two students valued questions that "expose you to new ideas that you'd never come across" and "writing exercises [which] expanded my view of writing quite a bit."

One of the reasons for the course being so engaging was attributed by approximately two in five students to the teacher. A diverse range of reasons were given of which the most common concerned the teacher's instructional behavior. One student relayed how her English teacher will "take a boring poem and point out things you'd never see and make it interesting." Other students described the way their teacher explained or structured the material, provided corrective feedback, and involved the students by inviting them to discuss their experiences in relation to the topic being studied. In the former case, for example, one student was impressed by the way his teacher "will start with like a piece of the question and work each piece together, and then come out with a broad answer, so that all of a sudden we will know it." Other teachers made a course interesting by telling jokes or having a sense of humor, relating interesting stories as examples, kidding around, and generally making class fun.

In addition to the relevance to daily life that made courses intrinsically interesting and worthwhile for many students, some students found a class engaging because of its relevance to their future goals. Statements such as "I'll use this knowledge after school," "in the future that's where the jobs are going to be," and "I

have to know it and I'll be doing more of it later" are indicative of this instrumental form of relevance. It is interesting to note that no students mentioned institutional rewards, such as good grades, as an explanation for the most engaging course, and only one described a reason related to instrumental performance (e.g., the material was easy to learn).

Table 6C below lists in rank order the most common categories of reasons for which (a) humanities (social studies and English), and (b) all other subjects, were described as the most interesting and worthwhile school subjects that students had taken in the year of this study.

Table 6C: Most Common Student Reasons for Nominating Humanities and Other Subjects as the Most Engaging Class

Rank	Humanities	Rank	All Other Subjects
1	intrinsic interest in subject matter	1	intrinsic interest in subject matter
2	class discussions of opinions/ideas	2	manipulating information or things
3	interesting/fun teacher	2	interesting/fun teacher

There is one notable difference between the reasons provided, on the one hand, for social studies and English classes being engaging, and on the other hand, for those mentioned in relation to other subjects (i.e., science, mathematics, foreign language, high technology lab, art, health, and photography). Although intrinsic interest in the subject matter was the most common reason in both cases, the second most frequent student explanation of engagement in humanities classes was the opportunity to discuss one's own ideas or opinions or (in two cases) to hear those of others. In contrast, this reason was never mentioned in the case of other subjects. Instead, one of the three most common reasons that non-humanities subjects are engaging is the opportunity to manipulate information or materials.

For most students the appeal of discussions appears to be either the chance to merely get involved in the lesson (and obtain some relief from "the teacher only talking which gets boring"), or being able to compare (or check out the status of) their own views with those of their peers. However, other reasons emerged, such as a genuine interest or curiosity in hearing different opinions, and the stimulation of the competition involved in trying to "win" an argument against one's peers. And a couple of students felt that discussions were more relaxing and easier because they demanded less effort than other types of lessons. The first explanation is analagous to actively participating in other subjects through problem-solving and "hands-on" activities.

Given previous findings that social studies is not perceived by students to be relevant to life (Shaughnessy & Haladyna, 1985), the high ranking of intrinsic interest - which half the students attributed to relevancy - is interesting. Again, it suggests that the social

studies teachers included in the HOT project sample do differ from typical teachers of this subject, at least in being able to make the subject matter intrinsically interesting.

Examples and Attributes of an Engaging Class

Each student was asked: Have you ever had a class or course where you found the subject so interesting that time passed much more quickly than usual and where you put forth your best effort? Every student answered in the affirmative. Then they were asked what was the subject of this class. The frequency with which each subject was mentioned is listed below in Table 6D.

Table 6D: Subject of Really Engaging Class (n=45)

Rank	Subject	Students Frequency	Percentage
1	Social Studies	14	31.1
2	Science	11	24.4
3	Mathematics	7	15.6
4	English	6	13.3
5	Foreign Language	3	6.7
6	Computer Science	2	4.4
6	Art	2	4.4
8	Technical Drawing	1	2.2
		46*	102.1*

* One student rated two subjects equally engaging.

In comparison with the most engaging courses taken during the year in which students were interviewed, social studies remained the most popular choice for a really engaging class ever taken in high school. It was closely followed by science which was nominated by approximately a quarter of the students. In fact, the number of students nominating science or mathematics (40%) was approximately the same as the number mentioning humanities subjects (i.e., social studies and English) (44.4%).

When asked to explain what made the class or course so interesting and why time passed so quickly, more than four out of five students gave a reason related to the intrinsic value of the subject matter or activities in the class (see Table 6E). The most frequent specific reason was again (as in the case of most engaging subject) that the subject matter was intrinsically interesting. More than half of the students who selected a social studies class mentioned this reason, such as one who found that "learning about the holocaust, the legal system, and the 1787 Constitution was really interesting, and SDI, Nicaragua and the contras - I really liked that." This student then revealed why she found these topics intrinsically interesting: "I wasn't sure about these things, like my mother and father talked about it [Nicaragua] a lot. [Now] I would talk to them about it and it was

really great [because] they'd bring up all these points." Other subject matter (in social studies in four cases and one each in biology, chemistry and computer science) was apparently intrinsically interesting because of its relevance to real world events or to a student's interests and concerns beyond school. Examples from social studies included: "some of the material was related to my father's law practice," world cultures is very interesting because "my mother is with the airlines and we travel a lot," and "she got us into it [genetics] because it was about us ... like the reasons why I have black hair or brown eyes."

Science was another subject that a number of students reported as intrinsically interesting: for example, "there's so many unbelievable things that just rattle my mind, like how little tiny organisms are found and how things are created."

Some students indicated that a class was particularly engaging because of the opportunity to manipulate information or materials. Nearly half these cases were mentioned in relation to mathematics, but this reason also was given for social studies, and "hands-on" activities in science and art. Two descriptions given for mathematics were "graphing a polynomial function is interesting because you can do so much with it," and "I like to figure out things where you have to work out or fix a problem." Another student described how in art class it "excited" her as the shape and form of a ceramic piece changed when she placed it in a kiln.

Over a third attributed the reason for a class being particularly interesting and worthwhile to the teacher. Presenting a variety of activities or topics and providing effective instruction (e.g., challenging and clearly structured tasks) were the most popular teacher behaviors, although nine different types were described. Other teacher behaviors mentioned by several students included giving us "fun activities" (examples of which are described in the next section),

Table 6E
Reasons a Particular Course or Class was Engaging
(n=45)

	no.*	%*	
instrumental performance in class 4.4%	1	2.2	material was easy to learn
	1	2.2	performed better than others
instrumental value of class	0		
intrinsic performance in class 2.2%	1	2.2	performed well
intrinsic value of class 82.2%	18	40.0	intrinsically interesting subject matter
	9	20.0	opportunity to manipulate information or materials
	6	13.3	cognitively challenging tasks
	6	13.3	opportunity for discussion or contribution of own ideas/opinions
	6	13.3	opportunity to create a unique or original product
	6	13.3	learned new things or gained new insights
	3	6.7	opportunity to hear ideas/opinions of others
	1	2.2	opportunity to revise previous material
teacher behaviors 37.8%	8	17.8	instructional practices: made subject interesting (fun activities, stories), challenging
	4	8.9	personal characteristics (likeable, enthusiastic, sense of humor)
	4	8.9	provided variety (of activities or topics)
	3	6.7	knowledgeable about subject
	3	6.7	attitude towards students (caring, gave responsibility)
peer behaviors 6.7%	2	4.4	students in class were all friends
	1	2.2	serious attitude to learning
	1	2.2	academically competitive

* Numbers add to more than 45 and percentages add to more than 100 as more than one reason could apply to any response.

being knowledgeable about the subject, and having a personable or likeable disposition.

An instrumental reason was given by only two students, and a reason associated with the behavior of their peers was mentioned by only four students.

Following the discussion of the particular class or course that each student selected as really engaging, a five-point scale was provided on which the student indicated how often he or she had been involved in such an engaging class throughout high school. Students were evenly split between the those who said fairly or very often, and those who reported only experiencing such engagement sometimes (the majority), once in a while or (for one student) practically never.

Most Engaging Activities or Tasks

Students were asked to describe some of the most interesting activities they did in the course where the subject was so interesting that time passed much more quickly than usual and where they put forth their best effort. Six students either could not recall a specific activity or indicated that there were no particularly engaging activities. When the descriptions of the remaining 39 students were categorized according to the four dimensional coding scheme (see Table 6F), there was only one dimension on which a majority described the same characteristic and that was on interaction where approximately sixty percent indicated an activity done alone. One-third of the students mentioned a writing task, but the remainder described seven of the eight other activities or tasks with only one, an experiment, being reported by more than ten percent.

Unfortunately, there was a high frequency (one in four) of uncodeable responses on the content dimension because many descriptions of engaging activities did not include details of the specific content and interviewers were not directed to probe for these details. Almost half of the codeable responses (or just over one third of all student responses) involved an activity or task in which the content was concerned with the principal ideas or central meaning of the material being studied. Some examples of such activities were:

-learning about symbolism and applying it to our own reading,

-doing research in the library for an essay on "The Federalist era: Did domestic and foreign policy endanger or secure the nation?" and then "seeing how historians view the same time period differently", and

-"The teacher demonstrated an experiment with a cartesian diver which sank when it was squeezed. We had to figure out why this

Table 6F
Frequencies and Percentages* of
Students' Descriptions of Activities in an Engaging Class

type	no.	percent
writing task	15	34.1
experiment	6	13.6
reading (only) task	4	9.1
discussion	4	9.1
debate, role-play	4	9.1
field trip	3	6.8
film	3	6.8
oral presentation	2	4.5
uncodeable	3	6.8
	44**	99.9

interaction	no.	percent
self	26	59.1
peers	10	22.7
teacher	5	11.4
uncodeable	3	6.8
	44**	100.0

content	no.	percent
ideas	17	35.4
experiences	8	16.7
values	4	8.3
specific topic	4	8.3
facts	2	4.2
uncodeable	13	27.0
	48***	99.9

function	no.	percent
manipulation	21	41.2
intake	13	25.5
creation	7	13.7
demonstration	3	5.9
uncodeable	7	13.7
	51***	100.0

* Of total number of responses on each dimension from 39 students who described an engaging lesson or assignment.
** Five students described two activities.
*** More than one category on this dimension could apply to any one response.

> happened. I spent hours and hours trying to figure it out, although it was only worth a couple of extra credit points for which you normally wouldn't spend too much time."

Some students described an activity concerned with understanding the experiences, feelings or emotions of oneself or others. Often this involved trying to understand the perspective of people who lived in a different time or culture, such as "movies which make a culture live by showing you the way of life and giving you a feel for the culture." And sometimes it involved journal writing "to express your views" on a painting, event or some other personal experience.

Given that almost one quarter of the students nominated their most engaging course (in the current school year) because of the opportunity it provided to contribute their opinions or ideas, it was somewhat surprising that less than 10 percent discussed an activity where value judgements had to be made or examined. Instead the activities described by these students ranged across all types of content.

Only two students mentioned an activity emphasizing the learning of discrete facts, but in both cases a higher level cognitive activity also was involved. For example, one student discussed doing a map of the eastern hemisphere on which "we had to name the countries", but added that they also were required to explain "why they had alliances with their different countries." This finding suggests that students do not find activities engaging when the content involves only isolated facts or fragmented bits of information.

In terms of the purpose of the interesting activities described by students, many of the codeable examples were intended apparently to have students manipulate information in order to make sense of it. Such activities included solving word problems and equations, organizing and synthesizing arguments for a debate on Indian land rights, analyzing information to cross-examine a witness in a simulated court trial, conducting a survey on an environmental problem, and formulating questions on an oceanography experiment for another group to answer. Activities concerned only with acquiring information represented one in four responses. Examples of these intake activities were: viewing slides on Spanish speaking countries, reading different books like All Quiet on the Western Front, listening to the teacher explain symbolism and to other students give their opinions on evolution, and going to an "art show for critical appreciation." A number of students mentioned activities that encouraged them to use newly acquired knowledge or skills to create an original product, such as developing a computer program, making a poster with a metaphorical representation of a concept, and writing original essays or reports.

Across the four dimensions, an individual writing task requiring the manipulation of ideas was the single most often described task, representing 20 percent of the responses. This finding contrasts with reports of teachers' complaints about student resistance to substantive academic tasks (Hampel, 1986; McNeil, 1986; Tye, 1985).

Engaging Social Studies Lessons and Assignments

Students were asked if they had experienced a lesson or an assignment in the social studies course observed by a researcher which they found so interesting that time passed much more quickly than usual, and where they made more effort than they normally do in school and really tried to do their best. All but eight students indicated that they had. These students then were asked to describe what the lesson or assignment involved.

The categorization of the responses on the four dimensional coding scheme (see Table 6G) resulted in a somewhat similar frequency configuration to that for the descriptions of activities in a particularly engaging class. A writing task again was the most often mentioned type of task (by almost a third), but in this instance was closely followed by discussion, with debate/role-play also reasonably popular. These last two activities generally occur more frequently in social studies (and English) than in other subjects and hence their higher frequencies in a social studies context are not surprising. Consistent with the popularity of writing assignments, individual tasks were more commonly described than activities or tasks involving interactions with peers or the teacher.

It is interesting to note that the vast majority of these individual tasks (or approximately one third of the engaging examples students than class activities, especially those enabling interactions with peers.

Ideas and values dominated the content of students' examples of engaging lessons or assignments in the HOT social studies course. Ideas were involved in such subject matter as the law of supply and demand, hypotheses about human behavior, symbols and images associated with Andrew Jackson's personification of American traits, the motives of the founding fathers (mentioned by three students), and interpretations of the Federalist era and the presidencies of George Washington and Thomas Jefferson. Value issues included such examples as deciding: whether the whiskey rebellion was justified, whether immigration was good for the

Table 6G
Frequencies and Percentages* for Students' Descriptions of Lessons and Tasks in Social Studies Course that are Engaging
(n=37)

type	no.	percent
writing task	13	31.7
discussion	10	24.4
debate, role-play	7	17.1
reading (only) task	2	4.9
questioning	2	4.9
oral presentation	2	4.9
film	2	4.9
uncodeable	3	7.3
	41**	100.1

interaction	no.	percent
self	17	41.5
peers	10	24.4
teacher	10	24.4
uncodeable	4	9.8
	41**	100.1

content	no.	percent
ideas	17	41.5
values	12	29.3
specific topic	5	12.2
experiences	2	4.9
facts	2	4.9
uncodeable	3	7.3
	41**	100.1

function	no.	percent
manipulation	22	53.7
intake	8	19.5
creation	8	19.5
demonstration	0	0
uncodeable	3	7.3
	41**	100.0

* Of total number of responses on each dimension from 37 students who described an engaging lesson or assignment.
** Four students described two activities from lesson/assignment.

country, what were the characteristics of a good congressman or senator, whether a nuclear freeze should be advocated (the example provided by all three students from one class), and whether there should be stricter gun control laws.

The function of the majority of lessons and tasks described by students was to manipulate or make sense of information, usually by analyzing ideas or evaluating issues such as those listed above. Writing a book critique, deciding one's position and developing supporting arguments for a class debate or essay, and constructing a visual and descriptive metaphor were examples of the kind of reported tasks whose primary purpose matched this category. In only one case was an example given of a task whose function essentially seemed to be enabling students to demonstrate the knowledge (or skills) they had acquired - and even here the student criticized that aspect of the task:

> "We did a written report on Jackson's personification of American traits. It was sort of stupid because all we did was put ideas in the book into our own words. We had to describe two or three specific images of America that Jackson was a symbol of. It was a 200 page book, but it didn't take that long to read because I found it really interesting. .....It was more interesting because you're dealing with symbols and images. One image was that Jackson thought that God was on his side. ....I did have to put the book in my own words, but the subject was more interesting."

These findings suggest that generally these students seem to be engaged by social studies tasks which allow them to actively explore various ideas and values. But to understand this preference more fully, we need to examine their reasons for finding these lessons or tasks engaging.

As Table 6H indicates, the vast majority of students reported that the lesson or assignment they had described was so interesting because

Table 6H
Reasons Social Studies Lesson or Task was Engaging
(n=35)

	no.*	%*	
instrumental value of task	0		
intrinsic performance on task 5.7%	2	5.7	performed well or did a good job
intrinsic value of task 82.9%	14	40.0	intrinsically interesting subject matter
	10	28.6	opportunity for discussion or contribution of own ideas/opinions
	9	25.7	learned new things or gained new insights
	5	14.3	cognitively challenging task
	5	14.3	opportunity to hear ideas/opinions of others
	3	8.6	opportunity to manipulate information or things
	2	5.7	made historical event or figure real
	2	5.7	opportunity to play unique role
teacher behaviors 14.3%	3	8.6	knowledgeable about subject
	2	5.7	instructional practices
	1	2.9	personal characteristics (likeable)
peer behaviors 2.9%	1	2.9	cooperative approach to work
uncodeable	1	2.9	

* Numbers add to more than 35 and percentages add to more than 100 as more than one reason could apply to any response.

of the intrinsic value of the task. Three specific reasons within this general category were frequently mentioned. The first, and most common, was that the subject matter was intrinsically interesting, which included such topics as the theory of natural selection, ancient Greek civilization, criminal trials, women's rights, and the Constitution. For some students it was a particular fact on one of these topics that was engaging: "I found it fascinating .. how we have been on earth for such a small fraction of the life of earth," and "you wanted to know the answer as to why this might be a fact because fact itself [i.e., people who drive white cars have a much higher rate of skin cancer] was so interesting." Other students, however, had a more complex explanation of why they found a topic intrinsically interesting, for example:

> "I liked writing about it [whether the founding fathers were trying to perpetuate their own class interests]. It's interesting because it goes back to Marx and class analysis. Taking economics got me interested in Marx and my father has been talking about Marx for years. This gave me a chance to apply critical views that are otherwise put down by students who have a Rambo-like mentality."

The above quotation also illustrates the second reported reason in the category of intrinsic task value, the opportunity to contribute one's own ideas. A role play of a debate at the Constitutional Convention provided such an opportunity for another student by enabling him, in his words, "to speak your mind and get across your ideas" through a particular character.

Learning something new or gaining a better insight represented the third common reason. Typical statements were: "I never realized before what actually happens" (in the process of arraignment), "I hadn't encountered it [the revisionist view] before, whereas everything else in American history I have had before," and:

> "In the first few days of this class the teacher put up the quotation: 'Where you stand depends on where you sit.' I found that interesting because it presents you with a different and more realistic side of American history. It was something new as I had never looked at it that way before."

Again these last two quotations indicate an interest in ideas or world views, rather than factual information (as in the first statement). Interestingly, no instrumental reasons were stated and only a few students directly attributed their engagement in a lesson or task to their teacher.

Engaging Questions or Tasks in Observed Social Studies Lessons

As well as the questions about school subjects in general and the social studies course in particular, students also were questioned on each of the three visits about a social studies lesson observed by a researcher. They were asked what, if anything, did they find interesting about that lesson and what made it interesting for them?

The responses were categorized on two dimensions: one involving the cognitive type of question or task, and the other the intended function or cognitive purpose (see Table 6I). In the latter case, half the responses could be categorized within the function of intake, that is they involved situations where the student was acquiring information, usually by listening to the teacher or to other students' answers to a question or task. The other half of the responses described questions or tasks where the student had to engage in active mental effort to make sense of information, or in a few instances create a unique product, in order to respond to a question or task. Some examples of manipulative function questions were: determining "why some leaders are better than others," arguing whether civil disobedience is justified, "trying to find out whether technology was making the government spend more or less - we really had to think about it," and "trying to list the similarities of the three industrial tycoons." An example of creation was a small group task requiring the creation of a court case involving First Amendment rights.

The most common cognitive type of question or task (described by almost half the students) concerned the relations between facts, such as comparisons and contrasts among different facts, and explanations

Table 6I
Description of Engaging Questions or Tasks in Observed Social Studies Lessons
(n=105)

cognitive type	no.	percent*
relations between facts	48	45.7
- required comprehension only	- 22	- 21.0
- required inference	- 15	- 14.3
- required abstraction	- 8	- 7.6
- required empathy	- 3	- 2.9
evaluation	24	22.8
- justifications of opinions	- 14	- 13.3
- opinions	- 10	- 9.5
specific facts or generalizations	22	21.0
definitions	1	1.0
uncodeable	10	9.5
	105	100.0
function		
intake	54	51.4
manipulation	46	43.8
creation	4	3.8
uncodeable	1	1.0
	105	100.0

* Percentages were computed by dividing the number of student responses in each category by the total number of responses (i.e., 105).

Note: 1. In 14 cases students indicated there were no engaging questions or tasks in an observed lesson.

2. In six cases students reported that the lesson in general was engaging, but could not think of a specific question or task.

3. Twelve students were not taking the same social studies course (i.e., the one included in the HOT sample) on the third visit and hence were questioned on only two observed lessons.

4. Three students were absent from the observed lessons on one visit and hence were questioned on only two lessons.

5. Five students described two engaging questions from an observed lesson.

(of events or states) giving causes, consequences or predictions (Hyman, 1979). If these comparisons and explanations were made for the student who merely acquired the information (that is, the function was intake), then the response was categorized as comprehension. Fourty percent of intake responses were comprehension of relations between facts. On the other hand, if the student actively participated in constructing an answer by interpreting, analyzing or manipulating information, then the description was categorized as demanding inference, abstraction, or empathy, using Newmann's (in press) delineations of content challenges in social studies. For example, the question (posed by a student) of "How did America avoid a revolution at the time, given the extremely bad working conditions and the pronounced separation between the classes in terms of wealth?" represented an inferential type. An example of abstraction was interpreting economic data and explaining how increasing the number of workers can eventually decrease revenue. And one of only three examples of empathy involved a role play of the Constitutional Convention in which a student had to try to understand and argue the position of her character in supporting both slavery and a Constitution intended to protect people's rights.

Evaluation questions or tasks (cited by almost one in four) involved either the justification of an opinion or value judgement, or merely the expression of an opinion or making of a value judgement. Explaining one's opinion on "what would you look for in a democratic leader?", and "defending the rights of the American Indians, that they had a moral argument to keep the land" were two responses in the first category. While most of the responses in the second category described making personal value judgements, a few responses emphasized hearing the opinions of peers and comparing them to the student's own views on a particular question. For example, one student stated her interest in seeing how the class classified different countries on various adjectives and in wondering "what other people are thinking of when they choose something different to you." Another student liked finding out people's views on, for example, South Africa so that he could "learn who I can talk with" and how "I can tailor my responses to get them to respond."

With only one exception, questions or tasks involving specific facts or generalizations involved an intake, or acquisition of information, function. A typical example was "finding out how we got our states of California and Texas - because America went to war for those states."

Questions intended to elicit an empirical response (that is, demanded facts, explanations or conclusions based on facts, or inferences drawn from facts) were cited in over 70 percent of the codeable responses. This result, however, may not necessarily represent the kinds of questions students find the most engaging, but reflect the dominant mode of questions that students face. Although we have no data on the respective percentages of question types posed during these observed lessons, a study of the teaching of a social studies topic

reported that over 80 percent of the questions were of the facts or relations between facts type (Bellack et al, 1966).

Engagement in In-depth Work

Concerns have been expressed about the predominant emphasis on coverage of content in high school classrooms and the negative efffect of this emphasis on student engagement (Newmann, 1988). Therefore, it seemed useful to find out whether or not students found the alternative to a coverage approach, that is in-depth work, engaging. Students were asked if they had ever had a class or course in which they spent a considerable amount of time (say, two or more weeks) digging deeply into a specific topic, question or problem. Almost two-thirds (64.4%) replied that they had. When asked in what class this experience occurred, nearly all responses were either social studies (14) or English (13), with science (4) being the only other subject mentioned by more than one student.

Those students who had experienced in-depth work were asked if there was anything they found interesting or worthwhile about this work. Only one student said there was nothing interesting or worthwhile, and three indicated that it depended on the topic or the amount of time (e.g., "beyond five days it either becomes boring or frustrating as you want to change your whole essay because you become confused and change your opinion on the essay question.").

Students generally stated that in-depth work was interesting because they liked the particular subject or topic, or because it provided the opportunity for active participation that enabled them to be creative or to work autonomously, such as doing their own research for a paper. For some students interest in or liking for the topic evolved from the process of in-depth study:

"If you take one subject that you don't understand at all, or don't think would interest you, and then over the course of two weeks, whatever, for me I just fall in love with that topic. ... It got to the point that this book [on the Red Baron fighter pilot] was part of my life...I wasn't nervous [in giving an oral report to the class] about anything, I didn't need notes, it was like all scarred in my head."

Similarly, other students indicated that their engagement resulted from active participation in the work. For example, one student reported "I had a lot of research to do. Doing things on my own ... going to the library and looking things up made it interesting for me", and another said "as I did the research I got more interested in it." Yet another student described a creative writing assignment in which her engagement stemmed from "coming up with an idea and a plot that I thought was great."

It was the worthwhileness of in-depth work that was emphasized by many students. As the following two examples illustrate, the result of

examining a topic in detail can be painless learning and enhanced understanding.

> "You learned a lot. We did a project on the praying mantis. We took pictures, went to the library, took lots of notes, made a billboard, and presented it to the class. ... You didn't realize that you were really learning. That's worthwhile!"

> "We spent three weeks devoted entirely to the constitutional convention. It was an incredible experience because you could really get the feelings of the events and share with the people what happened and why."

Probably the most compelling testimony for in-depth work, rather than an emphasis on coverage, was provided by a grade 11 student:

> "I got totally immersed in a project when the teacher forced us to do a paper on some guy. We couldn't pick him, but we had to read at least four books and write at least 100 note cards (big cards), and develop at least a 10 page paper. I got Montaigne. It ended up real interesting. As Mr. Dudley pointed out, it was kind of cool that I got to be a real expert and to know more about this guy than probably five million people in America. I'm not sure what made it so interesting - whether it was Montaigne's own works and life or just the fact that I got to know so much about him. Most of the time, you don't get this in school. A lot of times it's a total skim; it's very bad. A classic example is this course in European history. We covered 2000 years. Every week we had a 30 page chapter due. It's one of the hardest courses in the school. A real lot of work. He's a stickler for dates, all dates and the facts. We had 50 dates a week to memorize. The pity of it all is that now I don't remember anything. I worked so hard, and now basically all I remember is Montaigne. There's like maybe five dates I remember, when I probably learned three or four hundred dates all year. I can't even remember even a lot of the major guys we studied."

Despite a number of such powerful explanations of the value of in-depth work, the students were almost evenly divided between those who would like to have more of this type of work in school (15) and those who would not (12). Among those who did not want more in-depth work, several felt that it involved too much work even though it was worthwhile (e.g., "I'm lazy", "Because it was so hard"), while some preferred either to obtain a broad knowledge of a subject ("If you did that all year you wouldn't get to cover other topics.") or to have more variety ("I like variety with new stuff each day."). The student cited above who did the project on Montaigne wanted more in-depth study, but recognized the dilemma of concentrating on broad coverage or depth of knowledge:

> "Yes, I'd like to have more work where you dig in depth, but it's a double-edged sword, 'cause if you're constantly going in depth about each thing you come across, then you're not going to get very

far. It's quantity versus quality. The only reasonable thing is you've got to find a balance. I guess there's more of the superficial quantity in school now - teachers trying to cover as much as they can. They're not going really into depth."

Teacher Characteristics associated with Engagement

Having explained why a particular class was so interesting that time passed much more quickly than usual (and described some of the most engaging activities in that class), students also were asked what stands out in their mind about the teacher. Four students indicated nothing, but as Table 6J reveals, the rest described a diverse range of characteristics. Various instructional practices, particular attitudes towards students, and a number of personal characteristics were most often mentioned.

Making the subject interesting by providing fun or engaging activities, relating interesting stories or anecdotes, or in some other unspecified way was the most commonly described teacher behavior. One student said her Spanish teacher "teaches in a normal way but adds things that are fun: for example, playing games to learn or review for a test, singing songs, teaching the different meanings of certain words." Other teachers made the class interesting because they "made like the most boring stories exciting", such as one who "would really get into it ..he was like kind of a kid ...he would tell us all these interesting stories" and another who "remembers what he disliked about chemistry and tries to make that interesting by, for example, telling these bewildering facts: for example, about how one mole of rice would cover the entire world 60 meters deep, and then he explained how he got this figure."

A number of other general instructional practices made the teacher stand out for a few students in each case. These were: encouraging student participation in the class (e.g., "he gives students a lot of

Table 6J
Descriptions of Teachers in Engaging Class
(n=41)

	no.	%	
knowledge	5	12.2	knowledgeable about subject
17.1%	2	4.9	experientially knowledgeable about subject
	2	4.9	knowledgeable about adolescent learning needs and interests
	1	2.4	knowledgeable about a variety of topics or issues
instructional practices	13	31.7	makes subject interesting (e.g. fun activities, stories, anecdotes)
61.0%	4	9.8	encourages student participation
	4	9.8	challenges students' thinking
	3	7.3	gives lots of examples
	2	4.9	explains material at appropriate level
	2	4.9	gives well-organized lessons, clear tasks
	1	2.4	gives helpful feedback
attitude towards students	8	19.5	concerned about helping students understand the material
	7	17.1	gives students freedom/responsibility
46.3%	4	9.8	respects students' opinions, ideas
	3	7.3	fair, treats all students equally
	1	2.4	encourages and praises students' efforts
	1	2.4	shows genuine interest in adolescents
personal	8	19.5	sense of humor, tells jokes, etc.
character-	7	17.1	friendly, likeable person
istics	4	9.8	enthusiastic about teaching subject
43.9%	3	7.3	open-minded, accepts new ideas or presents all sides of an argument
	1	2.4	confident, articulate, warm speaking style
academic	2	4.9	strict demands for completing work
expectations	1	2.4	expects quality work, demands student's
7.3%			best effort

* Numbers add to more than 41 and percentages add to more than 100 as more than one category of teacher characteristics could apply to any response. Percentages were computed by dividing the number of students responding in each category by the total number of students (i.e., 41).

chance to talk, discuss and debate things"), challenging students' thinking (e.g., "he asks a lot of questions that you have to think about"), providing lots of examples ("to make it understandable"), explaining the material at the students' level (e.g., "she breaks down tough words into your language instead of stating them in scientific terms"), and giving well-organized lessons or providing clearly defined tasks (e.g., "she makes class flow easily - one thing ties into another, connections between units").

A group of attitudes towards students that could be characterized as reflecting a combination of a caring concern and an adult-like respect was mentioned by almost one in tworespondents. The most frequently discussed was a demonstrated concern for helping students to learn and understand the material being studied in class. Comments such as "she didn't get frustrated when I was getting the wrong answers and stuff and she helped me out a lot," and "she spends time on important things: asking about what you didn't understand" were typical of this category. Almost as many students identified their teacher's willingness to give them a certain amount of freedom and responsibility as a significant attribute. Examples students gave of this attitude were: "he trusts kids, for example with bunsen burners, and thinks we are responsible," "she allows a lot of freedom (for example, you can sit with your friends, you can bring radios into class)," and "she gives them a chance to do what they want." Showing respect for students' opinions and ideas (e.g., "she incorporated kids' ideas into the class," "he is polite and he accepts what you have to say even if sometimes you are wrong about something, he still respects your opinion") and treating all students fairly and equally (e.g., "he didn't favor any individual and gave everybody his time") were each reported by several students to be characteristics of their teacher that stood out.

Such personal characteristics as a sense of humor, a likeable or friendly personality, enthusiasm for teaching, and open-mindedness also were described by nearly one in two students. A teacher's willingness to occasionally be funny, tell jokes, or just generally exhibit a sense of humor was important to a number of students in making a class engaging, while being friendly or nice was engaging to others (e.g., "he keeps a friendly relationship with the students, all the class basically likes him"). A few students felt that being enthusiastic about and energetic in teaching (e.g., "she's into it, loves teaching") was an outstanding characteristic of their teacher, while some others designated having an open mind (manifested, for example, by a readiness to "accept new theories").

Only a relatively small number of students indicated that the teacher stood out because he or she was knowledgeable. Besides the predictable attribute of being knowledgeable about one's subject, which was mentioned surprisingly by only a handful, being knowledgeable about adolescents ("he knows what kids like and dislike") and about a variety of other subjects or issues ("you can bring out anything and she talk knowledgeably about it") were other categories described by a couple of

students. And only three students mentioned academic expectations as a feature of the teacher in their particularly engaging class that stood out in their mind. This last finding is particularly important given the emphasis in the effective schools literature on teachers' high expectations of all students (Purkey & Smith, 1983).

In addition to discussing the teacher of a particularly engaging class, students were questioned more generally about any teachers who stand out because they push them to put forth their best effort. Two questions were posed to the students (all but two) who indicated there were such teachers: How do they get the best from you? How do these teachers differ from others? Responses to these two questions were treated as one response and analyzed together since the second question essentially was asking students to elaborate on their first response by comparing the teacher(s) to others. Again instructional practices and attitudes towards students were the most frequently designated characteristics, but this time as many students mentioned academic expectations as personal characteristics (see Table 6K).

A concern for helping students understand the material was easily the most common response. Typical statements included "they'll give you help if you need it" and "they show they care if you learn," while "other teachers don't try to help you along with a problem, they just give it to you." A more unusual but perceptive comment was "they don't give you busy work, but just enough for you to understand the material." Some students also emphasized the teacher's patience and individual attention, as the following examples illustrate:

> "Well, he explains work to you and if you don't know and you try and do it and you still can't, he still helps you and doesn't get frustrated."
>
> "These teachers take time with you, they go over it again and again. They have patience."

Encouraging and praising students' efforts and showing a genuine

Table 6K
Descriptions of Teachers who Push Students to Best Effort
(n=43)

	no.*	%*	
knowledge	3	7.0	knowledgeable about subject content
7.0%	1	2.3	knowledgeable about adolescent needs and capacity for learning
instructional	8	18.6	encourages student participation
practices	8	18.6	makes subject interesting (e.g. fun activities, stories, anecdotes)
58.1%	6	14.0	challenges students' thinking
	4	9.3	explains material at appropriate level
	3	7.0	gives helpful feedback
	2	4.7	gives well-organized lessons, clear tasks
attitude	18	41.9	concerned about helping students understand the material
towards			
students	7	16.3	encourages and praises students' efforts
60.5%	5	11.6	shows genuine interest in and caring concern for adolescents
	4	9.3	gives students freedom/responsibility
	3	7.0	fair, treats all students equally
	3	7.0	respects students' opinions, ideas
personal	10	23.3	friendly, likeable person
character-	5	11.6	enthusiastic about teaching subject
istics	4	9.3	sense of humor, tells jokes, etc.
37.2%	1	2.3	confident, articulate, warm speaking style
academic	6	14.0	strict demands for completing work
expectations	5	11.6	assigns and requires a lot of work
39.5%	4	9.3	want you to work and do well without making strict demands
	1	2.3	expects quality work, demands student's best effort
	1	2.3	uses embarrassment or threats of punishment

* Numbers add to more than 43 and percentages add to more than 100 as more than one category of teacher characteristics could apply to any response. Percentages were computed by dividing the number of students responding in each category by the total number of students (i.e., 43).

interest in students' lives were two other ways in which identified teachers demonstrated a caring concern, according to a number of students. Examples of the former category included "she tells me that I can do it if I just keep trying" and "he keeps telling me how great the defense mechanisms assignment was, [and] makes you feel good about the way you think about something." Representative of the latter category were the comments: "she became a friend whom you can talk to about anything," "he knows you as an individual: if you come up to him after class he remembers about your personal life and past stories," but "other teachers don't seem to like kids or what they're doing: it's sad that schools don't look enough at whether teachers really like students." A willingness to give students a certain amount of freedom and responsibility, showing respect for students' opinions and ideas, and treating all students fairly and equally were other features of teachers' attitudes towards students that were important for three or four students in each case.

Instructional practices that commonly distinguished teachers who pushed students to their best effort were making the subject interesting, encouraging student participation, and challenging student thinking. While in the case of a particularly engaging class the first of these three was described by more than three times as many students as the other two, in this context it was mentioned less frequently with all three being about equally common. Examples of the first two practices already have been provided, while examples of challenging students to think were Socratic questioning (e.g., "they keep on asking you until you get the question - they really make you think"), and "creative" assignments that "really make you think about what you are doing."

Teachers' academic expectations featured more prominently in students' discussions here than in their descriptions of the engaging class. In addition to teachers who made students do a lot of work, 10 students in total characterized teachers who push them as either insisting that work be done (and usually on time) or conveying a similar expectation or desire without imposing strict demands through threats of punishment or other external means. Strict teachers, for example, "check the [home]work everyday", "get on your backs to meet deadlines", and are "pushing, pushing, pushing, ensuring you're working the whole time, watching over you in class." Other teachers, however, were reported to be more effective in eliciting other students' best effort by "sort of pressuring you ...[by] telling the consequences of what will happen, but not so much that you're stressed"; by "pushing you, but in a way that makes you want to do it"; or merely by indicating in some way that "they want you to work, but don't do anything specific to get you to work." This distinction in the responses of different students suggests that some students need external pressure to do their best, while others are motivated by more subtle means, such as the student who said of her teacher "if you don't do the work it feels like you're letting down a friend."

The three personal characteristics of a nice or likeable personality, a sense of humor, and enthusiasm for teaching again were the most common responses in this category. Almost one quarter of the students mentioned likeability which represented the second highest frequency of any category. In contrast, teachers' knowledge about their subject rated a mention by only three.

The low frequency of responses describing teachers' knowledge base as an outstanding characteristic (in classes which are engaging or those where students put forth their best effort) contrasts with the emphasis placed on this characteristic by the teaching effectiveness literature. It is possible, however, that the students in this study regard most of their teachers as knowledgeable about their subject and therefore do not find it a distinguishing characteristic.

Types of Lesson Formats that Students Find Engaging

Besides the open-ended questions on particular classes, students also were asked to indicate what types of lesson formats (in all subjects) they, in general, found interesting and worthwhile. Two out of three students agreed that student-centered class discussions (such as debates and role plays) are engaging (see Table 6L). And another three of the eight listed types were checked by a majority: working on a task in a small group, class discussion led by teacher, and working on a task with a friend. These four share a common instructional form, namely one in which students are active participants engaged in interacting with their peers (as well as in one case their teacher).

Class discussions in general (either teacher or student led) were explained to be engaging for a number of different reasons. Some students simply emphasized that they are "a good way to get everyone involved." But besides enabling active participation in class, others stated you are "exposed to different points of view." While this exposure for some may satisfy just an interest or curiosity in hearing different opinions, other students seemed to value the opportunity discussions provided for social comparison, that is to see how their own views compare with those of their peers:

> "You learn what the majority thinks. You hear other people's reactions to your points: whether or not they agree. ... I feel better when people have the same point of view as I do."

Table 6L: Types of Lessons Students Report as Engaging
(n=43)

Rank	Lesson Format	Frequency	Percentage
1	class discussion led by students	29	67.4
2	working on a task in a small group	24	55.8
3	class discussion led by teacher	23	53.5
4	working on a task with a friend	22	51.2
5	film	18	41.9
6	depends on the class or teacher	16	37.2
7	teacher questioning class on understanding of subject matter	15	34.9
8	working on a task by myself	11	25.6
9	lecture	10	23.3
10	other	1	2.3

When disagreements arose, a few students, in contrast, enjoyed the competitive challenge of trying to win an argument with their peers: for example, "if you really 'zing' someone it builds up your self-esteem." The challenge for some, however, was to make "you think about your own" [view] and "understand [it] - which you must do to explain it to others."

Students who preferred student-led discussions tended to stress that everyone was more willing "to say what's on their minds" because, according to one student, "you don't feel inferior as you do with a teacher," or because, according to another, "sometimes you are more willing to listen to your classmates" (than the teacher). As a result student-centered discussions were felt to produce more lively arguments. Two students also mentioned the benefits of debates where you may have "to be on a side you don't agree with and then you're well-informed about both sides of the topic."

The advantages of teacher-led discussions, on the other hand, were described as having the teacher add "correct information" or "facts and ideas students might not think of." Furthermore, the teacher "can lead it in interesting directions" and keep the discussion "orderly", focused and relevant which "avoids going off into things that don't matter to the lesson."

Working on tasks in a small group or with a friend was perceived as engaging for similar reasons. They provided "more relaxed" and "personal" environments in which it was "easier to express yourself" and "hear everyone's ideas" and "different opinions." Various students also claimed that in small groups you get "more ideas" and you "come up with the best answer by working together," and "if you don't understand you're more willing to admit you don't understand and can learn from the others." Some students seemed to find such conditions only came from working with a friend, while others only specifically described small groups in this way. Although many students said that it was easy to work with a friend because, for example, "you know how a friend

thinks," some admitted that "you can easily get distracted because you want to talk about other things."

One student made an interesting distinction between different subject areas in regard to the respective advantages of working with a friend and working alone. He reported coming up with the best answers in math and science by working alone because "you know what you did" in analyzing and trying to solve a problem, whereas in social studies and English he finds he does better by "talking back and forth" with a friend. But another student stated that working on a social studies or English task alone enables you to "do what you want to do."

Films typically were portrayed as "more relaxing and easier to understand" because they "usually explain what's going on more clearly," as well as just providing a change of pace. Teacher questioning was emphasized (by approximately one third) as helping "you understand something" (which "makes it more interesting") by getting "feedback from the teacher and the class," and making "you more comfortable about giving your own opinion" (after hearing the teacher's opinion or questions). Less than a quarter of the students reported lectures to be engaging, but those who did said they were worthwhile because, for example, "you have to pay attention" and "you can get a better understanding of what the teacher is talking about if they sit up and explain it to you step by step." Often, however, interest in lectures depended on the particular teachers, some of whom "are boring." And apparently "some teachers can even make discussions boring," but "mostly the topics are boring." Engagement in class discussions also for one student depended "a lot on the teacher because some teachers go too fast or expect you to know too much, while others ignore half the class."

Types of Writing Tasks Found Engaging

Students also were requested to designate the types of writing and the subjects of writing assignments that they generally found interesting and worthwhile. In the former case, writing paragraphs or a short essay, taking notes, and writing a paper were each checked by approximately half the students (see Table 6M). Almost as many, however, indicated that it depended on their interest in the topic or subject whether or not any type of writing was interesting. Interestingly, the less substantial writing tasks of making an outline and writing one or two sentences were checked as engaging by the fewest numbers of students.

Writing paragraphs or a short essay was reported to be interesting because "it gives you a chance to be creative and come up with your own point of view," as well as worthwhile because "you have to select the most important points and write clearly and concisely." Many students seemed to feel that short essays are a good length because they don't require an excessive amount of information, but allow sufficient scope for "spreading facts around and giving different opinions" in a concise and specific form. Other students, however, found that short essays

enable you to "only get out one or two ideas, [which] you can't really expand," whereas in a paper "I can get all my thoughts out." One student emphasized that writing a paper "puts you in charge of what you want to do and learn," adding that it is "a very personalized kind of task that you are responsible for."

Table 6M: Types of Writing Students Report as Engaging
(n=43)

Rank	Type of Writing	Frequency	Percentage
1	writing paragraphs or a short essay	22	51.2
2	taking notes (from a lecture or book)	20	46.5
3	writing a paper	19	44.2
4	depends on the topic or subject	18	41.9
5	making an outline	13	30.2
6	writing one or two sentence answers	4	9.3
7	other	3	7.0
8	none	1	2.3

Taking notes was stressed generally as being worthwhile rather than interesting. The reasons cited were the necessity for "concentrating" or having "to read the material carefully and understand it before you can re-word it," or the advantage afterwards of having the notes to refer back to (especially helpful in preparing for tests, or to enable you "to see the progression of ideas in a class"). One student tried to explain how both the process of taking notes and the end product helped you learn:

> "There is something about taking notes that helps you learn, I can't explain it. It helps you because you're taking it in and learning something new. Having background notes on topics you don't know anything about makes learning so much easier."

Making an outline was regarded as worthwhile for somewhat similar reasons as those given for taking notes. In other words, these students reported that actually making an outline helped them understand and organize the main points or ideas, while the completed outline "helps when taking a test and writing an essay." Typical also of the comments of several students was the claim that it is "easier than writing the paper itself, you deal with major ideas only and not specifics which take more effort."

In response to the list of general subjects of writing assignments, taking a position on an issue and defending it was the most popular, being checked by approximately two out of three students (see Table 6N). Describing personal experiences or feelings and explaining a concept, principle, theory or problem were each indicated as engaging by about half the respondents, with creative writing popular with just less than half.

Table 6N: Subject of Writing Assignments Students Report as Engaging (n=43)

Rank	Type of Writing Assignment	Frequency	Percentage
1	taking a position on an issue and defending it	28	65.1
2	describing my experiences or feelings	22	51.2
3	explaining a concept, principle, theory, problem or issue	21	48.9
4	creative writing (making up stories or poems)	19	44.2
5	reporting on independent research	13	30.2
6	summarizing material covered in class or readings	11	25.6
7	depends on the topic or subject	9	20.9
8	other	2	4.7
9	none	1	2.3

The challenge involved in arguing for and defending a position on an issue seemed to underlie most students' explanation of why this type of writing assignment was engaging. For example, students commented that "you have to take into account arguments of the opposition and counter them with your arguments," "it's your own point of view and you try to persuade the reader to come to your point," and "it's challenging and you have to stay on one side of the argument the whole time." And "if you can't give reasons why you feel strongly about your position, then you realize you don't think as strongly as you believed you did." The challenge was not limited for some to defending their own position, because "even if you don't believe in it [the position], you can put forth the effort and discover you can argue for something you don't believe in."

Both journal and creative writing were appealing to many students because of the freedom they provided. Students said they didn't have to worry about: "spelling and grammar," "using notes or an outline," "following an assigned topic," "sticking to facts", and "definite answers that the teacher is looking for." Instead, they could "just make things up," "write how you talk," and write "whatever comes into your head." Describing personal experiences or feelings enabled several students to express and reveal part of themselves, and one student to "get a load off your back and ... think about and deal with my personal problems." Another student provided a compelling account of how journal writing enhanced the teacher-student relationships in a class because of the teacher's interest in "our thoughts and feelings." Furthermore, she added that discussions were better because "people in the class opened up more," for example, "if talking about a novel, they'd be more likely to share their true feelings."

Explaining a concept, principle, theory, etc. in writing was reported to be worthwhile because it helped you understand the material. The one quarter of respondents who checked summarizing

material generally indicated that it was worthwhile for the sane reason: for example, it "helps you understand what someone is saying if you have to summarize it." Those who found reporting on independent research interesting enjoyed the opportunity to choose their own topic which often occurred with this type of assignment. The research for this writing was regarded by some students as worthwhile because "you learn a lot about the subject" and "you're sharpening your research skills which are important irrespective of which field you enter."

SUMMARY:

Social studies was the most frequently mentioned subject when the students interviewed in this study were asked about the most engaging course they had taken in the current school year and about a particularly engaging class they had experienced in high school. Science, English and mathematics were the only other subjects mentioned by more than a couple of students.

In probing students' reasons for being engaged by these classes, and by lessons or tasks in a social studies course observed by a researcher, the subject matter and/or the activities were mentioned by the vast majority in each case. Half of them described the subject matter as intrinsically interesting (especially in the case of social studies and often because of its perceived relevance to the real world), and half reported that the opportunity to actively participate in class (mainly through discussions in social studies and English, and through manipulating information or things in science, mathematics and other subjects) made the course or lesson engaging. In the case of the social studies course, many cited as engaging lessons where they were learning something new. Many students also attributed their engagement specifically to the teacher, particularly his or her ability to make the class interesting or challenging.

When asked to describe some of the most interesting activities in a particularly engaging course and the nature of a particularly engaging lesson or assignment in their social studies course, a writing task was the most popular type in both contexts (closely followed by discussions in social studies). The content involved in most students' examples concerned main ideas and either values (in the social studies course) or people's experiences (in the case of an engaging course), rather than discrete facts. Somewhat surprisingly many of the engaging examples were homework assignments. In both cases, but especially in the social studies course, most students outlined activities requiring them to manipulate or make sense of information, while a number described activities concerned with acquiring information. However, when questioned about specific social studies lessons observed by a researcher half the students mentioned as engaging questions or tasks involving the acquisition of information, with just under a half describing questions or tasks requiring cognitive manipulation (mainly making inferences or value judgements).

All but one of the students who had experienced in-depth work (i.e., studying a specific topic, question or problem for two or more weeks) found it interesting. Generally they felt it was interesting because they liked either the particular topic or the opportunity to be creative or to work on their own. Many also emphasized how much they learned from doing in-depth work. However, they were almost evenly divided on the question of whether they would like to have more in-depth work in school.

When asked about any outstanding characteristics of the teacher in a course they found particularly engaging, about a third of the students described the teacher's ability to make the subject interesting (for example, by relating interesting stories or anecdotes, or assigning fun activities) and a third mentioned various other instructional practices. Attitudes towards students (such as being concerned about helping students learn and giving them a certain degree of freedom and responsibility) and personal characteristics (such as a sense of humor and a likeable disposition) were each listed by nearly half the students. Instructional practices (including encouraging student participation) were again mentioned by most in relation to teachers who push students to their best effort, with the same number discussing attitudes towards students, especially a teacher's concern for helping them learn. While academic expectations were rarely mentioned in the case of an engaging course, this time many students made some reference to teachers' academic demands.

In terms of general types of lesson format that are interesting and worthwhile, class discussions led by students were the most popular, followed by small group work, class discussions led by the teacher and working with a friend. Class discussions seemed to be engaging for a number of different reasons, ranging from creating student participation to enabling social comparison with peers. Working in a small group or with a friend was regarded as personal and relaxing.

As far as writing tasks are concerned, writing a short essay, taking notes and writing a paper were each indicated as interesting and worthwhile by approximately half the students. The length of short essays was popular with some students, while others preferred the greater scope available in a paper for expanding on their ideas. Taking notes was strictly seen as worthwhile for utilitarian reasons, namely facilitating concentration on the material and serving as a reference when studying for tests. Two out of three students indicated that taking a position on an issue and defending it was an engaging kind of writing assignment because of the challenge involved. About fifty percent in each case thought journal, creative and expository writing (about a concept, principle or theory) was engaging. Journal and creative writing were appreciated for the freedom they provided from worrying about having the right information and answers, while expository writing was desribed as valuable for helping you to understand the material.

OVERALL CONCLUSIONS ON COGNITIVE CHALLENGE AND ENGAGEMENT:

Many previous studies have indicated that high school students generally find classes boring and put minimal effort into their academic activities (Csikszentmihalti & Larson, 1984; Cusick, 1973; Goodlad, 1984; Sedlak et al, 1986). One of these studies (Csikszentmihalti & Larson, 1984), however, had found that students, on average, had at least one class which was engaging. Similarly, in this study the students were able to discuss at least one class and one academic task where time had passed quickly and they had made their best effort. Furthermore, their engaging academic experiences had similar attributes; that is, most students tended to be engaged by similar kinds of academic work and teacher behaviors. Students indicated that activities such as discussions and writing tasks that require active mental effort and teachers who make the subject interesting and convey a caring concern for their learning are likely to engage them. In particular, instructional formats which involve students in interacting with their peers (e.g., debates, role plays and small group work) or the teacher (e.g., class discussions), and writing tasks which require students to take and defend a position on an issue are appealing to the majority of students.

Beyond the descriptive characteristics of engaging academic work, the more fundamental question is: What are the underlying causes of student engagement? Despite popular conceptions, extrinsic rewards (such as high grades, college admission or future job prospects) were reported by only a few students as a factor in creating engagement. Instead, the intrinsic value of learning was cited consistently across different academic contexts by a substantial majority. Students valued academic learning because the content was interesting, the process was actively involving, or the outcomes were worthwhile. The content was usually interesting because it was relevant to concrete human events or to a student's idiosyncratic interests and concerns outside school. Outcomes were worthwhile if the student learned something new or developed a better understanding of the material being studied.

Different students were intellectually challenged by different kinds of school work. Essentially the students in this study were equally divided between those who found humanities subjects (i.e, social studies and English) challenging and those who were more challenged by science and mathematics. In many cases it was the different types of problem associated with each of these subject areas that distinguished students' perception of challenge. The former group tended to report that ill-defined problems with problematic solutions made them think hard, while the latter group often found well-defined problems with single, correct answers more mentally demanding. Another distinction was between the demands for inductive inferencing in humanities subjects cited by some students and the demands for understanding abstract concepts in science and mathematics mentioned by others. In short, students perceive the major challenges in English and especially social studies to involve the making of value judgements and

(usually inductive) inferences; and the major challenges in science and math to be making deductions and abstractions.

As far as specific activities and assignments are concerned, teacher-directed activites that required student responses (e.g., teacher-led class discussions and teacher questioning students on understanding of subject matter) and substantive writing tasks that were completed individually outside class were most commonly described as mentally demanding, especially when they dealt with ideas and/or values. In the latter case, the noteable example was papers or essays, which for some students were challenging when they involved expository writing and for others when expressive writing was required.

Again the important question to consider is: What makes these kinds of academic work cognitively challenging? Demands for higher order cognitive processing featured most prominently when students were asked if they experienced any frustrations. But these frustrations were apparently overcome since the most common satisfactions were from successfully accomplishing the task and from developing new insights or better understandings. This suggests that a challenge must be of sufficient difficulty to create an initial frustration, but not so difficult that it cannot be met - presumably with reasonable effort.

Although teachers have claimed that secondary students passively resist intellectually challenging tasks (Hampel, 1986; McNeil, 1986; Tye, 1985), little evidence was found to support this contention. On the contrary, the kinds of academic work most students reported as mentally challenging were also generally the kinds they found engaging. Essay writing tasks and teacher-led discussions dealing with ideas and values were frequently cited as both engaging and challenging (although student-led discussions and working with peers were engaging but not challenging). Additionally, many students are both engaged and challenged by classes that emphasize higher order thinking.

References

Bellack, A., Kliebard, H. M., Hyman, R. T., & Smith, F. L. (1966). The language of the classroom. New York: Teachers College Press.

Boyer, E. L. (1983). High school: A report on secondary education. New York: Harper & Row.

Csikszentmihalyi, M., & Larson, R. (1984). Being adolescent: Conflict and growth in the teenage years. New York: Basic Books.

Cusick, P. A. (1973). Inside high school: The students' world. New York: Holt, Rinehart & Winston.

Cusick, P. A. (1983). The egalitarian ideal and the American high school: Studies of three schools. New York: Longman.

Everhart, R. B. (1983). Reading, writing and resistance: Adolescence and labor in a junior high school. Boston: Routledge & Kegan Paul.

Goodlad, J. I. (1984). A place called school: Prospects for the future. New York: McGraw-Hill.

Hampel, R. L. (1986). The last little citadel: American high schools since 1940. Boston: Houghton Mifflin.

Hyman, R. T. (1979). Strategic questioning. Englewood Cliffs, NJ: Prentice-Hall.

McNeil, L. M. (1983). Defensive teaching and classroom control. In M. W. Apple & J. L. Weis (Eds.), Ideology and practice in schools (pp. 114-143). Philadelphia, PA: Temple University Press.

McNeil, L. M. (1986). Contradictions of control: School structure and school knowledge. New York: Routledge & Kegan Paul.

Mosher, R., & MacGowan, B. (1985). Assessing student engagement in secondary schools: Alternative conceptions, strategies of assessing, and instruments. Paper prepared for Wisconsin Center for Education Research, Madison, WI.

Newmann, F. M. (1988). Can depth replace coverage in the high school curriculum? Phi Delta Kappan, 69(5), 345-348.

Newmann, F. M. (in press). Higher order thinking in the teaching of social studies: Connections between theory and practice. In D. Perkins, J. Segal, & J. Voss (Eds.), Informal reasoning and education. Hillsdale, NJ: Erlbaum.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). The shopping mall high school: Winners and losers in the educational marketplace. Boston: Houghton Mifflin.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. The Elementary School Journal, 83(4), 427-452.

Rutter, M., Maughan, B., Mortimer, P., & Ouston, J. (1979). Fifteen thousand hours: Secondary schools and their effects on children. Cambridge, MA: Harvard University Press.

Sedlak, M. W., Wheeler, C. W., Pullin, D. C., & Cusick, P. A. (1986). Selling students short: Classroom bargains and academic reform in the American high school. New York: Teachers College Press.

Sizer, T. S. (1984). Horace's compromise: The dilemma of the American high school. Boston: Houghton Mifflin.

Shaughnessy, J. M., & Haladyna, T. M. (1985, November/December). Research on student attitudes toward social studies. Social Education, pp. 692-695.

Tye, B. B. (1985). Multiple realities: A study of 13 American high schools. Lanham, MD: University Press of America.

Wehlage, G., Stone, C., & Kliebard, H. (1980, October). Dropouts and schools: Case studies of the dilemmas educators face. Madison, WI: Report to the Governor's Employment and Training Office.

Wynne, E. (1980). Looking at schools: Good, bad, and indifferent. Lexington, MA: Lexington Books.

VII

DEPARTMENTAL AND PRINCIPAL LEADERSHIP IN PROMOTING HIGHER ORDER THINKING

Cameron McCarthy
Francis Schrag

Many factors beyond the classroom may affect the quality of thinking in social studies classrooms. Do the principal and department head have a pronounced influence on what happens at the classroom level? Extensive questionnaires and interviews with administrators (as well as teachers and students) in the schools we studied, give us an opportunity to address this question.[1] Here we report some of the results of our investigation into the role of the department heads and principals in the five schools. We are still in the midst of completing our analysis of the impact of these actors on teacher's perceptions of support for higher order thinking and on what actually occurs in the classroom. At this point, however, we shall concentrate on profiles of the leaders at the department and school levels and identification of some similarities and differences between them. Since we are able to give a rough indication of the relative success of the five schools in promoting higher order thinking in social studies classes, we conclude with some speculations regarding the impact of the principal and department head on the classrooms.

I.

Department Chairs

A convenient place to start is with the department head who, in most high schools, plays an important role in hiring teachers, in assignment and development of courses, in selection of texts and examinations, and in setting the tone for the social studies program and department. We shall try to identify the nature and extent of support that each of our five department heads gave to their teachers in promoting thoughtfulness.

The five schools are very different in a number of important respects--geographical location, size of departments, number of periods in the day, average class size, and so on--but perhaps the most striking thing about the five social studies department chairs is the extent to which they agree on the principal satisfactions and limitations of their role. Almost all find their greatest satisfaction in watching successful teachers who have profited from their advice and support. On the other hand, almost all find dealing with recalcitrant

---

[1] Respondents may strive to put themselves and their institutions in a favorable light. But our assurances that the names of people and institutions would be disguised in the final report gave those we interviewed the opportunity to speak with considerable candor.

teachers who are not doing a satisfactory job to be their major source of frustration. The administrative side of the job--responsibility for meetings, paperwork, budgets, etc. is viewed by most as a necessary evil; only one mentions it as a source of gratification.

Department heads are also veteran teachers, and see themselves as peers. Although administratively superior, they have little actual power over their colleagues. Hence, their primary source of influence lies in their ability to win colleagues over to their side, to persuade and cajole rather than command them to go along with a particular program or point of view. It is no wonder that all five mention skill in interpersonal relations as the principal requisite for a successful department head. Neither is it surprising that they see themselves as most effective in informal, dyadic or small-group interactions rather than in staff meetings.

Despite commonalities which result primarily from the nature of the department chair's role and working environment, differences among the five department chairs we interviewed can be readily discerned. We shall briefly describe each of them.

Bob Mikkelson's title is "district coordinator" for social studies in Grandville, a wealthy, suburban school district in the Northeast region of the country. He has been in this position for 10 years and has been teaching social studies for 18 years. Under his leadership, the district's high school with 2000 students, (indeed the entire K-12 system) has been involved for the last 3-4 years in the development of a program to promote higher order thinking in the social studies area. This is by far the most systematic and continuous effort among the five schools in this area. Mikkelson is the only department head that one could really call the <u>architect</u> of a program.

Program development included the following activities:

1. Soliciting the assistance of a number of consultants, of whom the most influential has been Barry Beyer, whose explicit "skills" model of thinking has been adopted by the department.
2. Visiting other schools in the district and conducting a needs assessment within the school.
3. Designing an overall curriculum content and skills sequence for K-12, and revising a mandatory U.S. History course to incorporate the "thinking skills" approach. Formulating lesson plans and teaching demonstration lessons to staff.
4. Organizing systematic peer observation throughout the department.
5. Ordering materials and designing criterion-referenced tests to reflect the new focus.

Though much of this work was done by committees under Mikkelson's leadership during the regular school year, it is important to note that

the work continued during the summer months. Teachers were paid $75.00 a day for participating in developing the program; Mikkelson calls this the "most valuable" money.

Like all department heads, Mikkelson is responsible for observing teachers. He is required to make at least three visits to each of 40 district social studies teachers in a given year. To be able to carry on a vast program of curriculum development, while performing his other supervisory duties and teaching a class of his own, clearly requires unusual energy and dedication. Mikkelson believes that all his staff share his vision of excellence, though a few are not sure they have the tools to implement it. Mikkelson notes that he has helped shape and lead to fruition a major program change, and that he can go into classrooms and "see it being done." This he feels is the most rewarding part of his job.

Sam Holmberg is a veteran of 34 years teaching social studies. For the last 17 of these, he has been department chair at Carlsberg High School, a school of almost 3000 students in a major city in the Northeast. While Sam Holmberg is also enormously skilled and energetic, his approach differs from Bob Mikkelson in two significant respects. (1) Although he and his teachers have been involved in the revision of particular courses within the program, there is no explicit emphasis on thinking, per se. Carlsberg doesn't distinguish between instruction in thinking and good social studies instruction which "calls for thought-provoking questions and how and why questions." (2) While Mikkelson's emphasis is on the broad outlines and articulation of the elements of the entire social-studies curriculum, Holmberg's focus is on the individual lesson. If Bob Mikkelson is an architect, Sam Holmberg is a mentor.

The improvement of lessons is worked at from several directions. Holmberg does a great deal of observation, visiting each teacher at least twice a year, and non-tenured teachers at least six times a year. He writes detailed reports and conducts follow-up conferences with teachers concerning observed strengths and weaknesses. He, other teachers, even the principal on occasion, teach "demonstration lessons" to regular classes with colleagues in attendance. After the lesson, there is a critique session in which strengths and weaknesses are discussed. But the designing of the lesson plan receives greatest attention. Each lesson plan is a product of group effort, and every lesson plan taught is available to teachers in a resource file. Courses are perceived as sets of self-contained "lessons", so a major route to course improvement is via improved lesson plans.

Sam Holmberg, like Bob Mikkelson but in a less formal way, creates a climate in which sharing and collaboration are the norm. Holmberg sees his role as disseminator of critical ideas and teaching techniques. Asked about the most rewarding part of his job, he says, "Watching teachers perfect their teaching techniques is by far most important. It has been a tremendous pleasure to see Matt O'Brien become the great teacher he is. The specific pleasure comes from the

exchange of ideas and the planning that is an important part of teaching."

Herb Dexter, another veteran teacher of unusual energy, is department chair at Arnold High, a 1100 student high school in a middle-class suburb of a major West Coast city. While teaching a full load of courses (five per semester in a six period day), Dexter has pioneered the development of an eclectic approach to teaching thinking which he calls the "Integrative Mind Instructional Model." A striking feature of the model, developed collaboratively with one of his teachers, is the emphasis on metaphors and analogies which students create and represent visually in different colors. Like Mikkelson at Grandville, Dexter has read the conceptual literature on thinking, and has himself given workshops at other schools on his own approach to promoting thinking. He's also been involved at the state level in the creation of a new American History exam. Because the exam will emphasize thinking through issues, Dexter does not perceive it as a threat to his own emphasis in social studies.

Dexter sees himself and tries to act as a catalyst for others, to prod, stimulate, and support their efforts to excel as teachers. Because he teaches a full load (He was teaching 174 students each day in five classes when we interviewed him) he has neither the time nor the mandate to formally observe other teachers. Considering the limited free periods a day (each teacher has one), the enormous student load each teacher carries, the lack of formal supervision or peer observation, and the paucity of supplies, Dexter's ability to disseminate his vision and his ideas is remarkable. He claims that 85-90% of the teachers now share his general orientation.

Herb Dexter tries to share ideas with teachers in informal interactions, in curriculum planning groups which he organizes, and at formal department meetings which typically focus on pedagogical issues. He encourages others to take risks in their own teaching by revealing his own attempts to innovate, even when they're not successful. He treats everyone, even student teachers, as peers from whom he can learn something. The most important quality a department chair needs is "the intangible of inspiring people to keep shooting for the moon--trying to improve." Speaking of himself, Dexter says, "I am a supporter of others' rise to excellence--the role I cherish and sometimes abhor when my own intellectual commitment to a professional identity as a historian or political scientist emerges."

Louise Ellsworth, an especially thoughtful teacher, is the only woman in this group. Bradley High School, where she is a chair of social studies, is a 2100 student school in an older suburb of a large city on the Eastern seaboard. Its students display a diversity of background and ability not found in most suburbs. Ellsworth has been chair only three years. Prior to that, there has been a history of turnover in the role, due, she believes, to a tradition of individual faculty autonomy at the school which undermines the chair's ability to set and carry out policy at a departmental level. Ellsworth has a

general vision of what an ideal program to develop thoughtfulness would look like, but she is not naive enough to believe that most of the other teachers are looking for someone to lead them. One of Ellsworth's highest priorities is simply to develop a more collaborative culture in the department. A new required course for sophomores will provide an opportunity for teachers to work together.

The school system does not encourage systematic program development or in-service activities to improve pedagogy. In 1984, the department had some workshops with leaders of the "thinking skills" movement, but Ellsworth says only a few teachers really showed an interest. To the extent that the teachers do emphasize thinking in their classes, this results from their own personal orientation rather than from any department-wide efforts to change curriculum or pedagogy.

Ellsworth is required to confer with both non-tenured and tenured teachers several times a year, and these sessions provide the main official vehicle for working with teachers. Ellsworth believes that discussions about curriculum, approaches with students, or evaluation can be even more valuable than observing classes. "How much can you really tell a teacher about how to improve little moves in the classroom? Since most of the teacher's work is done before or after class, this is what demands more attention."

Unlike most of the other department chairs, Ellsworth has had the opportunity to hire a significant number of new teachers recently--eight. With young teachers, she feels her main task is to be supportive, to convince them that "every lesson doesn't have to be perfect." With older teachers who are comfortable with one style of pedagogy, her main challenge is to help them remain open to alternative approaches. "I encouraged a fairly rigid teacher who loves the Socratic method to use journals for kids to reflect on their own thinking."

We choose the label <u>therapist</u> to characterize Louise Ellsworth since much of her success with her staff as well as her orientation to teaching derives from her intuitive understanding of sound therapeutic principles: the need to build trust so that others can open up and share their shortcomings, the need to overcome resistances to new approaches and perspectives, the realization that change results from people being able to take a fresh look at themselves, not from supervisors pointing out their mistakes to them. Not surprisingly, she feels the least rewarding part of her job is "Working with teachers who are afraid to look at themselves."

Mike Preminger, the final member of our group, is social studies chair at Scarborough High, an 1000 student inner-city high school located in a major Eastern metropolis, a magnet school which plays a key role in a system-wide reform effort. A general focus on developing critical thinking for all students has been introduced at the district level, but it is institutionalized primarily through a set of writing

exercises and short discussions that are not systematically integrated into the curriculum of any department. Preminger was one of ten teachers from the district who served on a planning committee in development of a US History course.

In the social studies department, all three mandatory year-long courses are required to follow unit packets designed by a team of two university professors and twenty teachers. This can substantially undercut the autonomy of the individual teachers.

Almost all the social studies teachers, are veterans and were specially selected for this flagship school. Despite the attempt to develop a district-wide, "syllabus-driven" curriculum, the teachers value their autonomy and, like those at Bradley, do not appear to be looking for guidance or leadership. "The lone-ranger syndrome can hurt us," Preminger says. Although required to observe and confer with his teachers, Preminger feels this activity may not have much impact. Although he relished the chance to observe and give useful, positive feedback to a friend and colleague, Preminger feels he can best assist his staff by scheduling classes according to the strengths and interests of the teachers, by securing materials for use in their classrooms and by clearing other bureaucratic hurdles.

Unlike Bradley and Arnold, where there is little pressure from above, here at Scarborough there are constant demands and expectations from the principal and district. In the interviews, Preminger constantly referred to what "they" want. His main frustration is that although he was picked, in part, to develop a new interdisciplinary curriculum in international studies, the effort has never received the leadership from the school administration needed to bring it to fruition. Yet, of the five department heads, Preminger is the only one who claims that administrative activity is satisfying. "The materials and budgeting are very important. It gives me a better sense of control--especially the chance to find interesting materials that I can pass on to colleagues." His role appears to be primarily that of <u>intermediary</u>, relaying mandates from above, and requests from below. His leadership in curriculum development or promoting new approaches to teaching appears very limited, but it is hard to determine whether this is due primarily to his non-directive approach or to demands from above which seem to afford little opportunity for autonomous action at the departmental level.

Having briefly described each of the chairs and departments, we may observe that several have common features which not all share. In Grandville, Arnold, and Carlsberg there is a real sense of collegiality, of individuals not only working towards common goals, but also sharing ideas and helping each other perform better. The focus is on instruction, and the department heads in each of these schools are clearly perceived by the teachers as master teachers from whom they can learn. In Bradley and Scarborough, on the other hand, there is less sense of a common enterprise. Most of the veteran teachers are not looking for guidance from peers or superiors.

Another pattern is revealing. At three of the schools, Arnold, Carlsberg, and Grandville, their social studies department had a systematic[2] program aimed at promoting higher order thinking. In only one case, Grandville, had they adopted the "thinking skills" orientation. Scarborough, though it had school and district-wide programs focused on thinking, had no program that originated in the social studies department. Neither Scarborough nor Bradley manifested an explicit concern with higher order thinking at the department level.

Finally, it is worth noting, that only at Grandville, was there a major commitment of resources, both time and material, to both develop and implement a wholesale curriculum reform. Inadequate time for planning, too heavy a teaching load, lack of supplies of resources, failing equipment, seem to be common in the other schools.

II.

Principals

As we found in the case of the department heads, there are commonalities in the principals' conceptualizations and efforts in support of higher order thinking in the five schools we studied. But as we shall see, there are important differences as well--particularly with respect to administrative style and type of support provided for thinking. These differences in styles and approaches also influence the quality of collaboration and cooperation between the principals and the department heads. There is, for instance, a high degree of continuity and collaboration on agendas for instructional improvement and methods of evaluation between the principal and department head at Carlsberg. On the other hand, there is significant discontinuity between the perspectives and agendas of the principal and department head at Scarborough.

But first, let us discuss some of the commonalities among the principals. Even though there is little significant exposure to recent research literature, there is a remarkably shared vision among the principals of what a higher order thinking classroom should look like.

---

[2]According to our definition, a systematic departmental-based program aimed toward higher order thinking should reflect three criteria: members of the department should share a common conception and vision of higher order thinking; lesson plans, syllabi and other materials should relate the conception or vision to the lessons taught; and there should be interaction among department members on a regular basis that is focused on how well they are progressing toward the vision. Departmental programs can conceivably be grounded in district level or school-wide programs, but unless we found indicators such as these at the department level, we did not recognize the existence of a systematic, departmental-based effort.

This vision involves: less teacher dominated classrooms; teachers posing challenging questions or tasks; teachers allowing students to tackle questions from diverse perspectives. Four of the five principals report that their respective social studies departments are ahead of other subject departments in terms of classroom instruction in thinking. Two principals felt that a number of social studies teachers at their schools were not good "thinkers" and consequently were not good models of thoughtfulness for the students to emulate. But in neither of these cases, did the principals mentioned report direct involvement or intervention on their part to help improve the classroom instruction of the teachers identified as inadequate. Day-to-day instructional improvement in these two schools is seen by the principals as the responsibility of the department chairs.

All five principals attach an enormous significance to school "tone" and the fostering of an academically disciplined environment as preconditions for thinking ("we need an orderly and safe school concomitant with academic achievement," Lester Brown, Carlsberg). It is, however, significant that when asked about their goals, none of the principals lists thinking as his first priority. The three most consistently mentioned goals are: (1) improving school climate; (2) improving/maintaining academic achievement as reflected in test scores; and (3) teacher training.

Barriers to thinking -- The principals also agree about what are the main barriers and constraints on their efforts to promote thinking. Like the department heads, the principals report that administrative demands ("paper work") have been increasing over the years. These demands limit the amount of time they have to address curriculum and instructional issues in their schools. Three of the five principals (Arnold, Carlsberg, & Bradley) complain of scarce resources and lack of funds for instructional improvement and staff development programs in higher order thinking. At all five schools, curricular materials for higher order thinking, particularly teacher produced materials are considered scarce or inadequate. The principals also reported pockets of apathy and resistance to initiatives to higher order thinking among members of their staff. Though the principals readily express frustrations with scarcity of resources and funds, and constraints on their time to address instructional improvement, principals rarely mentioned organizational and structural features of their schools as significant constraints on higher order thinking. For the most part, the principals seem to accept as given the routine features of school organization, viz., class size, scheduling, and structure of the subject disciplines. They do not emphasize organizational or structural reform as a strategy for achieving higher order thinking in their schools.

Administration versus instructional leadership -- Even more so than the department heads, the principals experience a critical role dilemma -- a dilemma that is inscribed in the designation of school principal. This dilemma is expressed in terms of a tension between the demands placed on them as administrators versus demands placed on them

to be assertive curriculum and instructional leaders. Paul Benton at Arnold puts the matter rather forthrightly: "Administrative reorganization in the district has shifted my time away from curriculum issues to administrative work." Paul Benton's social studies department head, Herb Dexter, also mentions excessive administrative demands as a barrier to promoting higher order thinking instruction.

But it would be wrong to conclude from the above observations that the principals are not involved in instructional improvement. All five principals express a particularly strong interest in instructional issues and indicate that they make efforts to reconcile the dual obligations of administration and instruction. The principals all report that they use administrative controls over budgets, resources, personnel selection, agenda setting, and evaluation to influence instructional objectives in social studies and other subject departments at their schools. Three principals, Quentin Donato at Grandville, Ken Olds at Scarborough, and Martin Broderick at Bradley indicate that they use primarily administrative leverage, particularly, over personnel selection, to secure curriculum and instructional objectives. The other two principals, Lester Brown at Carlsberg and Paul Benton at Arnold, in addition to carrying out administrative duties, are directly involved in the fostering of instructional improvement and staff development programs at their schools. Unlike their counterparts at Bradley, Scarborough and Grandville, the principals of Carlsberg and Arnold see themselves, first and foremost, as instructional leaders.

Differences in Styles and Approaches to the Fostering of Thinking

Leadership styles and approaches to promoting thinking vary significantly among the five principals. These styles and approaches to thinking are influenced by such factors as school history/tradition, district policies, department head and principal relations, and the general climate within the school. We illustrate this variability by briefly describing the context within which each of the five principals operates and the salient aspects of each principal's leadership style and method of supporting thinking in school.

Lester Brown at Carlsberg works in a school environment that can be described as remarkably stable--this despite the fact that the school population is highly diverse--50% white, 22% black, 21% Asian and 7% hispanic. As was alluded to earlier, there is a very low rate of turnover among the faculty at Carlsberg. This is especially so in the social studies department where the average number of years teaching at Carlsberg, among our sample, is over 15 years. On the other hand, Carlsberg does not have the level of monetary and technical support for instructional improvement from its school district as we will see in the case of Scarborough and Grandville. Indeed, Carlsberg's social studies department head, Sam Holmberg, complains about the inadequacy of funds and resources for the promotion of higher order thinking instruction in his department.

What stands out about Carlsberg, however, is its strong sense of commitment to student achievement and academic excellence. The sense of commitment is shared by both the staff and students. Ninety percent of the students at Carlsberg go on to college. Students excel in district and statewide exams and competitions for scholarships and awards. This emphasis on academics is the Carlsberg "tradition," the principal suggests over and over again. He notes with pride, "we are an academic high school. If you come to Carlsberg, you'll achieve and go to college. All students will get an academic core curriculum." Brown takes direct responsibility for instructional leadership at Carlsberg.

Like his four counterparts, the principal at Carlsberg employs general administrative strategies such as personnel selection and budget control to secure his curriculum and instructional agenda. However, Lester Brown's main strategy for promoting higher order thinking is to be distinguished in one particular respect: Carlsberg's principal places enormous emphasis on thorough and meticulous evaluation and instructional supervision of his staff. In this respect, his collaboration and close working relationship with his department heads, particularly, his social studies department head, Holmberg, is crucial. Lester Brown holds workshops with his department heads "to help improve and refine their methods of observation and assessment." He does over 100 observations (including elaborately written summaries/critiques) of his teaching staff every year--more than any other principal in our sample. In addition, he reads and comments upon the lesson observation written up by his department heads. Lester Brown collaborates with Sam Holmberg in setting instructional goals for the social studies department and insists that these goals are met. He also pays meticulous attention to the details of pedagogical practices, exam results, and achievement levels.

Besides this emphasis on evaluation and his regular collaboration with the social studies department head, Lester Brown's image as an instructional leader is reinforced by his direct participation in and sometimes initiation of instructional improvement programs at Carlsberg. Lester Brown occasionally teaches in the classroom, gives demonstration lessons and leads workshops on instruction and pedagogy.

If Lester Brown at Carlsberg pays careful attention to the details of instructional supervision and evaluation, Scarborough's Ken Olds focuses more on administrative matters and coordination of district initiatives on thinking at Scarborough and elsewhere in the Scarborough school system. Furthermore, while his counterparts at Arnold, Carlsberg, Bradley, and Grandville are assigned to their respective buildings fulltime, Ken Olds is 50% time at Scarborough. He spends the other 50% of his time on his responsibilities to the district-wide training program on management skills for principals and assistant principals in the school district. Ken Olds actively works at two jobs--one in which he carries district initiatives to several schools and the other as the building principal at Scarborough. Olds' role is further complicated by the fact that Scarborough is a predominantly

black school which is currently under court order to integrate. The school has attempted to fulfill this requirement without using the expedient of enforced busing. It has therefore created magnet programs, such as International Studies and High Tech, to attract high achieving students who are mostly white. In the last four years or so, the percentage of white students in the school increased from 10% to 32% (the percentage of black students is now 66%). Ken Olds role as principal and the type of support he is able to provide for thinking at Scarborough must therefore be seen in the context of the school's desegregation efforts.

Unlike Carlsberg, where district monetary and technical support for thinking is regarded as inadequate, Scarborough is the beneficiary of relatively substantial district support for higher order thinking. The district has in this sense, sought to promote this inner-city school (only 38% of Scarborough's students go on to college) as a model institution. Ken Olds describes the relationship between the district and school in the following terms: "the central office (the district) has a major impact on the school's curriculum and instructional practices." This "major impact" is reflected in the fact that Scarborough has the largest number of higher order thinking related programs and the highest institutional profile for higher order thinking of the five schools studied.

Central to the instructional innovations being piloted at the school is the Scarborough Districtwide Critical Thinking Program (SDCTP). The program is funded jointly by the district and a large corporate benefactor to the tune of $300,000 over a two year period. The program involves 44% of Scarborough faculty. Teacher training and staff development in this program focus on curriculum and instructional strategies for promoting students "higher level thought processes." Emphasis is placed on divergent thinking and open-ended classroom discussion. To this end, a special interdisciplinary reader series, Response, has been developed for use in the classroom. The reader consists of short, often provocative passages on a variety of topics in philosophy, aesthetics, history, etc., which are intended to move teachers away from the traditional role of content dissemination and to stimulate student centered discussions. Another vital component of the SDCTP is the Syllabus Driven Exam an innovative approach to evaluation in social studies and other disciplines that aligns district tests with district curriculum syllabi. The Syllabus Driven Exam places emphasis on extended essay writing and is an attempt to provide an alternative to standardized or multiple choice tests.

Programs such as SDCTP, magnet programs in arts and sciences, and a teacher center (established for district-wide teacher training) help to underscore Scarborough's status as a model public school in the district. But this status is a mixed blessing for the school and the principal. The complicated internal demands of running a multiplicity of innovative programs along with his responsibilities to the district has meant that Ken Olds has had to delegate much of his responsibility for day-to-day running of the school to his assistant principals and

other members of staff. These substantial external and administrative demands significantly reduce the amount of time the principal can devote to curriculum and instruction matters at Scarborough. Ken Olds indicates that his involvement in instructional supervision has declined. "I have the additional responsibility this year of working as a resource to principals in other schools. Fifty percent of my time is spent away from the school. As a result, my classroom observations have dropped considerably."

While Ken Olds' involvement in the day-to-day instructional matters at Scarborough is limited, he reports that he uses his administrative leverage in the school to influence desired instructional goals. This is done primarily by means of manipulation of resources ("inducements"). In addition, Olds "encourages" his teaching staff to be self reliant. He espouses a philosophy of "empowerment" and believes that decision making over curriculum and instructional issues should happen at the department level. But Scarborough's social studies department head, Mike Preminger, does not regard himself as an instructional leader. He defines his role within the department as that of an intermediary who gets his colleagues the curriculum materials and other resources they need for classroom instruction. In some respects then, the philosophies of Scarborough's principal and social studies department head contradict rather than complement each other. While Olds claims to rely upon department-based instructional initiative, Preminger suggests that instructional leadership should come from the top. These different perspectives seem to have led to the fact that initiative and control over the higher order thinking instructional program have been assumed by the district. Rather than generating its own instructional agenda, the social studies department at Scarborough follows the district's approach to higher order thinking.

District leadership for Scarborough has gone further than any of the other schools in our sample to institutionalize instructional programs for the promotion of higher order thinking. But the externally motivated institutional innovations seem initially to have diminished teacher control over the form and pacing of the instructional program. However, Preminger expressed some hope that following the pilot phase for the syllalbus driven courses, the nature of exam questions would be sufficiently specified and limited by the district so that teachers would have more time and autonomy to teach non-examined topics.

For Martin Broderick, issues of higher order thinking, indeed issues of instructional improvement in general, have been secondary to an overwhelming concern with bringing "order" and "civility" to Bradley. Unlike, Scarborough there are no district motivated or school directed programs to promote higher order thinking instruction. The general approach with respect to instructional issues is to rely upon department chairs to fine-tune the curriculum, but according to both the principal and the social studies department head, teachers at Bradley have a lot of autonomy. Although Broderick has some ideas and

vague plans for higher order thinking, he had not developed them into programs. According to the principal, it is the department heads, like Louise Ellsworth in social studies, who should define the curriculum and instructional agenda. Martin Broderick provides the administrative supports and inducements:

> The primary instructional leadership responsibilities rest at department head level, department heads define the agenda for the departments...the principal supports these efforts by recommending fiscal support through workshop monies and "days away" from school, etc. [Martin Broderick.]

Broderick's emphasis on administration and order has to do, he says, with the "near chaos" he inherited when he arrived at Bradley about seven years ago:

> When I first came to Bradley, the school had a number of serious problems: overt and covert racism, violence, vandalism, graffiti, low achievement of blacks, low teacher morale--much tension, kids leaving to attend private schools. The main challenge was to bring order and stability. (M.B.)

The next important issue for Broderick was the curriculum. At the time of his arrival, liberal education at Bradley was a "completely elective curriculum". Although he believes in faculty responsibility for the curriculum, he has been concerned with the fragmentation and chaos spawned by too many electives. Therefore, Martin Broderick has insisted on some core course requirements for graduation and he sees his contribution to the curriculum as primarily one of imposing "structure" and "focus." Although he urged a tightening of the graduation requirements, he did not immerse himself in specific issues of content within courses.

Martin Broderick believes that the issues of school climate and the laissez-faire curriculum must both be addressed before the school can embark fully on a substantial program in instructional innovation. For Broderick, Bradley is at a transitional stage with respect to higher order thinking, but he sees major opportunities for this beyond the classroom in the new system of school governance he initiated. In the "town meeting," the students and faculty participate in decision making on a variety of issues such as class attendance, graduation requirements, and student behavior (eating, use of tape players, etc). For Broderick, the town meeting is an extracurricular experience in higher order thinking that allows students to participate in the governance of the school. The town meeting is considered by the principal to be a forum in which the students can challenge each other and also their teachers over important ideas. However, Broderick is less involved in pressing for higher order thinking in the classroom portion of the instructional program at Bradley. Like the principal at Scarborough, Martin Broderick adopts an indirect approach to

instructional improvement in his school. He is far less involved in supervision and evaluation of his teaching staff than principals in the other schools, each of whom report visiting twice as many social studies teachers per semester.

Despite his relatively limited contact with Bradley's teachers, Broderick supports Ellsworth's efforts in social studies, and he ranks the social studies department among the top departments at the school. Ellsworth, in turn, sees the principal as providing primarily moral support, rather than specific intellectual leadership, for higher order thinking initiatives at Bradley. Broderick views his contribution as providing the prerequisites for higher order thinking, namely, in bringing a much needed coherence and sense of focus to the curriculum and the liberal education program at Bradley; and in setting a school tone ("an orderly, disciplined academic atmosphere") for thinking.

Like Broderick, the principal at Grandville, Quentin Donato, also leaves the task of close instructional supervision to his department heads ("I do few observations"). However, he makes an impression as a systematic and methodical thinker--with a firm and clear position on the kind of thinking that he believes must be promoted in the school. Grandville's principal offers a more sophisticated conception of higher order thinking than any of the other principals -- a conception which involves a critique of Bloom's taxonomy ("Bloom's taxonomy is, I think, useful but only as a descriptive not prescriptive model."). Quentin Donato reports that he systematically supports thinking in the school and the social studies department through the following methods: (a) promoting peer supervision; (b) personnel selection and recruitment; (c) monies and resources for staff development; and (d) encouraging an interdisciplinary focus.

Two factors at Grandville reinforce Donato's initiatives. First, the school climate is stable, and enriched by a highly committed faculty and a tradition of academic achievement. Over 70% of Grandville's students go to college--almost twice the percentage of students at Scarborough. Secondly, the school has a unique relationship with the district office in that the district's social studies coordinator is based at the school and leads the crusade for higher order thinking.

Grandville's sense of institutional stability results in part from the relatively low rate of turnover of faculty and academic personnel. They have had only three principals in the last 50 years. Donato notes that there is a "common college-oriented set of values among the students, staff, and parents." The principal has made a concerted effort to promote peer supervision and a sense of collegiality and community among faculty and students. For example, members of faculty are strongly encouraged to work with student organizations, and there is a student-faculty council at Grandville somewhat similar in spirit to the town meeting idea at Bradley. Donato believes that extracurricular activities such as the student-faculty council help to promote a climate of receptivity to thinking among students.

The principal gets solid support from the district office. Indeed, the district office has established a definition of critical thinking and has initiated efforts through its subject coordinators to implement its agenda for curriculum and instruction in the schools. Grandville benefits from having such a subject coordinator, Bob Mikkelson, on staff in its social studies department. As we noted earlier, Grandville has developed a systematic approach to thinking that is teacher produced and central to its instructional program. Donato and Mikkelson collaborate in promoting the critical thinking program at Grandville. The principal, for instance, indicates that the realization of the critical thinking program in social studies at Grandville would have been "inconceivable" without Mikkelson. He has operated as both the architect as well as the chief practitioner of higher order thinking instruction in the social studies department. The district has provided significant improvement monies ($17,000) per year that provide support for outside consultants, staff development programs, peer observation, summer curriculum writing, and workshops related to higher order thinking. Grandville's teachers have also been able to visit and learn from colleagues in other schools.

It would seem that more than any school in our sample, Grandville benefits from a fortuitous relationship with the district which is both supportive but allows for a degree of autonomy. This type of district support and the collaborative efforts of the principal and department head have helped to sell the Grandville critical thinking program to the teachers.

Arnold does not enjoy the kind of direct district support that is found at Grandville and Scarborough. Both the principal, Paul Benton, and the department head, Herb Dexter, at Arnold criticize the district for failing to establish an initiative with respect to critical thinking. Paul Benton feels that the district has its priorities somewhat misplaced:

> There is no significant contribution or promotion [of higher order thinking] from the district. District concerns are more mundane, more focused on test scoresand developing a consistent and district-wide curriculumat this point. Aside from funding one or twomentorships...There has been no direct district interest or awareness [P.B.].

Paul Benton also complains that "administrative reorganization" in the district has significantly reduced the amount of time that he has to spend on curriculum and instructional improvement ("Too much paper work").

Instructional momentum on thinking at Arnold is directly due to the support of the principal, and the unflagging efforts of his social studies department head, Dexter, and individual mentor teachers such as Bill Strayhorn. The relationship between Arnold's principal and the

social studies department head is one of trust and mutual respect. Paul Benton feels comfortable in allowing Dexter and the social studies department to have considerable autonomy. He does not feel that it is necessary for him to be directly involved in the social studies department "because the department is by far the best department in the school." Dexter, in turn, regards Benton's moral support and guidance as crucial to instructional initiatives in higher order thinking at Arnold.

Arnold's principal sees himself as a curriculum and instructional leader..."something of a task master." He takes pride in this image of himself. "I work very hard and I expect my staff to work hard...I expect performance and results." Paul Benton, like Lester Brown at Carlsberg, is an instructional leader as well as an active administrator who is keenly in touch with the daily operations of his school. He spends a lot of time talking and working with teachers to get a better understanding of their needs. He pays special attention to new staff members who "need help in the classroom." Paul Benton shares responsibility for promoting effective teaching with his department heads and eight senior/mentor teachers at the school. These mentor teachers are funded by the district. At Arnold, they work closely with other members of the staff to help improve the overall quality of instruction. With the principal's urging, mentor teachers also hold workshops and implement innovative curriculum projects. Arnold's writing project, led by Bill Strayhorn which includes a strong thinking component and involves an interdisciplinary collaboration between the English and Social Studies departments is a prime example of the mentor initiative.

In addition to his active promotion of instructional improvement, Paul Benton deliberately uses personnel selection, recruitment, and cultivation of leadership to achieve his goals. According to Benton, when he arrived at Arnold about six years ago, "the average staff member was fifty-four, tired and scaling down." Benton set about revitalizing the staff, and within the last three years, there has been a 30% turnover. Benton set about revitalizing the staff. He used "every opportunity to hire quality teachers interested in instructional improvement and taking on challenges in the classroom."

Closely related to his emphasis on personnel selection, the principal has worked actively to "cultivate instructional leaders" among Arnold's teachers and youth leaders among the student body. Benton regards this "quality circle" of teachers and students as "partners" in his efforts to develop a momentum for thinking at Arnold.

III.

Conclusions

What, if anything, can we say about the impact of the department head's and the principal's leadership on the level of thinking actually found in the classroom?

Let us begin with a necessarily cursory overview of our findings at the classroom level. A total of 165 social studies classes were observed by one or two observers. Each class was rated along 14 different 5-point scales, of which six were picked as constitutive of Higher Order Thinking. The numbers below in parentheses represent an average of these six different scales. Five is the highest possible score and one, the lowest.

The highest scoring school was Carlsberg (4.13). Next came Grandville (3.99), then Arnold (3.91), Bradley (3.50) and, finally Scarborough (3.15). These ratings are not simply reflections of the individual teachers in the five schools. Differences between schools account for 28% of the total variance. The highest school differs from the lowest by 1.70 standard deviations, and from the fourth-place school by .93 standard deviations. Thus, institutional effects are important. Furthermore, the scores are not just reflections of the socioeconomic or racial composition of the five student bodies. For example, only half the student body at Carlsberg is white compared to 75% at Bradley. Let us now consider a number of plausible propositions about institutional effects and see how our data speak to them.

1. A systematic, department-based program, focused on higher-order thinking is necessary for obtaining classrooms which score high in this area.

The evidence from our study supports this. The three higher-scoring schools stand out in this respect although there are differences in the way in which they concentrate on higher-order thinking. Both Arnold and Grandville try to transmit a particular conception of HOT, whereas Carlberg considers the stimulation of student thinking as just one facet of pedagogical excellence. Scarborough, although it has a school-wide program, has no systematic effort at the departmental level. Finally, Bradley has no systematic program in the area.

2. Success in this enterprise requires substantial resources, including personnel, set aside for this purpose.

This is not supported. Grandville has, for example, allocated substantial resources into its higher order thinking program while Arnold has virtually none, yet their scores are about the same. Of course, the fact that Arnold can succeed with limited resources provides no guarantee that other schools can.

3. Success in this area depends on strong leadership at the department level.

This notion receives strong support. The department heads at Carlsberg, Grandville, and Arnold are dynamic leaders who inspire commitment. They work energetically to improve the quality of thinking

in their department's classes. The department heads at Bradley and Scarborough differ in many ways, but both take a relatively non-directive stance when it comes to shaping their colleagues' teaching styles in particular directions.

4. The principal must play an active role in supporting the department head's efforts.

This notion receives strong support. In all three of the top schools, principals were very active in working closely with department heads, but in neither of the bottom two did this occur.

5. District-wide efforts will be in vain if the teachers and their immediate supervisors, the department chairs, do not feel "ownership" of the program.

This notion, derived from so much recent literature on the subject (e.g., Stevenson, 1987), is clearly supported in the two schools that have district-wide programs. Grandville and Scarborough present the clearest contrast. In one case (Grandville), the initiative has come from the person who was also the department head in a single high school district; in the other case, the department head of a multi-school district appears to be left out of any of the school-wide efforts.

6. Success depends upon an institutional culture of collegiality which involves consistent, focused discussion of teaching and curriculum within the department.

This is confirmed: the top three but not the bottom two schools may be accurately characterized in this way.

Our study reinforces what many others have found, that the institutional context does make a difference. Collegial school culture and pedagogical leadership from department heads and principals help teachers promote higher order thinking in their classrooms. There do not appear to be simple, administrative actions which can be counted on to produce the culture that raises the level of thinking in classrooms. Whatever actions may be needed, our study indicates that energetic instructional leadership at the building level plays a significant role.

Reference

Stevenson, R.B. (1987). Staff development for effective secondary schools: A synthesis of research. Madison, WI: National Center on Effective Secondary Schools.

VIII

CONCLUSION

Fred M. Newmann

In spite of numerous proposals that thinking ought to be emphasized more in schools, studies heretofore have paid little attention to such basic questions as

> To what extent can higher order thinking actually find a prominent place in high schools that are conventionally organized?
>
> Why is it so difficult to promote?
>
> Why have some schools made successful movement toward this goal, and what have they done to overcome some of the barriers?

Answering these questions involves a number of steps, some of which have been completed in this study, but others await future research. Rather than summarizing each of the papers separately, I will attempt to synthesize the findings as they pertain to these issues.

We have made progress on the first question by developing a conception of higher order thinking applicable to teaching a variety of content and by operationalizing this conception into a reliable scheme for classroom observation. By studying a group of five departments selected for their special efforts to promote higher order thinking, we may have also obtained estimates of the "upper limits" of thoughtfulness in conventionally organized social studies departments. Whether the levels of discourse we found are truly exemplary will not be determined until we study a more representative set of departments in the next phase of the research, but at this point we have evidence that some departments demonstrate in their lessons high levels on the following dimensions:

> sustained examination of a few topics;
> substantive coherence and continuity;
> sufficient time for students to respond;
> the posing of challenging, non-routine intellectual tasks;
> teacher modeling of thoughtful problem solving behavior;
> students offering explanations for their conclusions.

In contrast, over all lessons we found less evidence of several dimensions of thoughtfulness: teachers' careful consideration of explanations and reasons, Socratic questioning, integrating students' personal experience into the lesson, encouraging students to come up with original ideas, questioning authoritative sources. Although such dimensions were relatively rare across all lessons, the highest scoring lessons, when compared with the lowest, did manifest more teacher careful consideration of reasons and Socratic questioning, and their students more often generated original ideas and articulate, relevant comments. These classes also relied more on class discussion, rather

than lecture, and the use of sources other than textbooks. While there was considerable variation in thoughtfulness both within and between schools, we found some impressive examples of students being challenged to use their minds.

In addition to researchers' observations of cognitive challenge in classes, the vast majority of students told us about being challenged and engaged in certain classes. Students nominated social studies more frequently than any other subject as the most challenging class taken in the present school year and also throughout high school. And most students indicated that the class that was most challenging also was their most engaging class. When asked to describe the kind of social studies work they considered most challenging, students identified such tasks as forming one's opinion, making inferences, and other tasks consistent with our definition of higher order challenge.

Teachers offered information on barriers to promoting higher order thinking. Across all schools, teachers believe that large class size, large numbers of classes per day, and large student load restrict their opportunity to promote thinking. There is also wide agreement that the pressure to cover content, whether externally- or self-imposed, often inhibits thinking. Teachers occasionally complained about lack of appropriate instructional materials and mandated tests, but these were not prominent barriers. Teachers in all the schools generally found adequate collegial and administrative support for an emphasis on thinking.

Student ability or background did not appear as a barrier to classroom thoughtfulness in these schools. In the interviews teachers rarely attributed problems in promoting thinking to student characteristics. Quantitative analysis of lesson scores showed only minor effects of student background, and the teachers who scored the highest taught more low achieving students than did their lower scoring colleagues.

Comparisons between the high and lower scoring teachers suggested that important barriers may reside in teachers' goals and conceptions of their work. Perhaps not surprisingly, teachers with the most thoughtful classes placed higher priority on thinking as an educational goal, articulated more elaborate conceptions of thinking, and placed more value on depth than breadth of coverage. This raises, but cannot answer, the question of the extent to which teacher beliefs must be considered the central target for reform in this area, in contrast, for example, to materials and organizational conditions.

How might we explain differences in classroom thoughtfulness between schools? While most variation in thoughtfulness is due to differences between teachers, more than 25 percent of the variation can be attributed to schools. The five schools differ in important ways, and the small sample will not allow powerful generalization, but we found support for some propositions that distinguish the top three from the bottom two schools.

Each of the top three schools, but neither of the bottom two, had some systematic, departmentally-based program that supported higher order thinking. The content of the "programs" differed substantially: one followed a nationally promoted scheme of convergent thinking skills; another emphasized a standard lesson format that began with a lesson aim stated as a problem or question and moved from informational to evaluative levels of analysis; the third placed major emphasis on divergent thinking and visual representation of thinking through metaphors and diagrams. The lowest scoring school was part of an extensive district-wide program, but here there was little sense of departmental ownership. The top three schools also manifested a departmental culture of collegial planning absent in the bottom two, and they similarly distinguished themselves through vigorous instructional leadership aimed at higher order thinking by both the department chair and the principal.

Together these findings indicate that classroom thoughtfulness in social studies can be assessed and that it can occur at high levels among a critical mass of teachers in conventionally organized high schools and with classes composed of students of all grade and achievement levels. We have also found that students consider thoughtful classes to be the most challenging and engaging, and we have identified the particular attributes that students mention in making such judgments.

In spite of this good news, we find what appear to be important differences in thinking between teachers who promote higher order thinking most and least consistently. Although we cannot point to particular administrative actions which guarantee the promotion of thinking in social studies departments, it seems that departmentally-based programs, collegial faculty culture and strong leadership by the chair and principal may be critical.

Still, there is much to learn. We have not investigated the relationship between classroom thoughtfulness and individual student achievement. Neither have we tested the extent to which these five select departments actually promote higher levels of thoughtfulness than a more representative set. Finally, although we may find high levels of thoughtfulness in some conventionally organized high schools, concerns about organizational barriers expressed by teachers and reformers invites inquiry into a third set of departments - those with structures modified more deliberately to promote thoughtfulness. During the next two years, each of these problems will be studied.